

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## Evaluating Evaluations of Clinical Decision Support Systems: Case Studies From NHS Clinical Settings

### Thesis

#### How to cite:

Dune, Richard (2016). Evaluating Evaluations of Clinical Decision Support Systems: Case Studies From NHS Clinical Settings. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2016 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

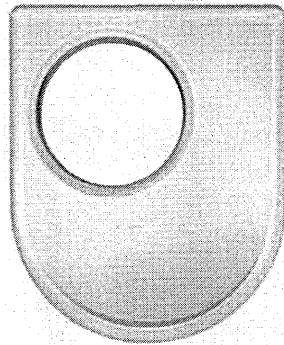
<http://dx.doi.org/doi:10.21954/ou.ro.0000efb1>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)



The Open University

---

# Evaluating evaluations of clinical decision support systems: case studies from NHS clinical settings

---

Richard Dune  
*BSc, MA*

A thesis submitted towards fulfilment of the requirements for the degree of  
*Doctor of Philosophy*

Department of Engineering and Innovation  
Faculty of Mathematics, Computing and Technology  
The Open University  
Milton Keynes, United Kingdom

---

SEPTEMBER 30, 2015  
THE OPEN UNIVERSITY  
Walton Hall, Milton Keynes, MK7 9AA

DATE OF SUBMISSION: 30 SEPTEMBER 2015  
DATE OF AWARD: 23 APRIL 2016

ProQuest Number: 13834825

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13834825

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

I dedicate this work to my parents, Elisha and Stella Dune. It has been a true labour of love and I am grateful for your love and support with every endeavour that I have taken so far.

To my life partner Elsie and my son Jaden Tatenda, thank you for all the support and putting up with me over the past few years. Jaden, we can go to the park a lot more now and also read some fun books for a change.

Also to all my siblings - Ruth, Rose, Rennie, Timothy and Rita; my nephews and nieces – Tanaka, Nyasha, Noku, Ceddy, Tashinga, Kundai and Kudzai; and my brothers in law – Leonard and Gilbert - I missed you guys so much and will certainly make up for the lost time.

My dearest childhood friends - Tendekai (it's never too late); Arthur (this was all your fault); Revai (it's always great to catch up with you) and Tapson (RIP) – life wouldn't been the same without our shared experiences.

To Reason (brother from another mother) and GoddyP (I still don't understand how we got through all that stuff!), thank you guys for the support and helping me make Coventry 'home'.

Steve Chindove and Mike Macheka, thank you guys for always being there for me.

---



## **Acknowledgements**

I would like to thank my supervisors, Professor Joyce Fortune and Dr Clive Savory for giving this opportunity and completely transforming me. I am truly grateful for the boundless support and direction, and also for putting up with me. I will forever treasure the past few years.

Donna Deacon, thank you for literally walking me through my PhD from the first to the last day. You are a superstar! Also thank you to the Open University for giving me this wonderful opportunity to learn and grow in this esteemed institution.

I would also like to thank all the people I have worked with throughout my clinical career in the NHS, which certainly has been a great help with this project (sadly I cannot mention any names for ethical reasons).

To Professor Jeremy Wyatt, Sylvia Wyatt, Dr Eberhard Mandishona and Dr Yollanda Chikari, thank you very much for sharing your insights at the very early stages of this project. I would also like to thank my “mini viva” examiners, Professor Geoff Peters and Dr Alison Bettley for your advice and input. Also, a big thank you to Professor Harry Boer (for the long chats about theories) and all the support from the Continuous Innovation Network members and fellow PhD students. To Dr Giles Brindley, Taddeus Mbaluku and Reason Ndlovu, thank you for your support and insights. To my colleagues at Regal Training & co., Lewis Normoyle, Jasmin Bhayat and Josh Lewis, thank you for your support.

## **Abstract**

The NHS is under increasing pressure to cut costs while delivering high quality care. At the same time, the demand for healthcare services has grown, driven in part by the increasing number of older people in the population. NHS Trusts are adopting clinical decision support systems (CDSSs) to help decision making at the point of care. CDSSs are said to bring benefits such as improvements in guideline adherence, clinical processes and user performance but evidence of these benefits is not always available and their effectiveness in terms of improving patient outcomes is often open to question. This thesis presents research that was carried out in a large teaching NHS Trust looking at the evaluations of three CDSSs. Semi structured interviews were carried out with key informants who were involved in their adoption, use and evaluations. Documentary analysis and observations were also used to augment the interviews. Most evaluations were carried out informally by the developers and were primarily driven by external regulatory pressures rather than patient outcomes and organisational needs. Evaluation documentation was inadequate or missing, thus making it difficult to systematically assess these evaluations. This thesis contends that evaluations are important to provide decision makers in NHS Trusts with adequate information to make decisions about CDSSs and computerised healthcare information technologies in general. NHS Trusts need to build organisational capacity and readiness to enable them to effectively carry out evaluations that will provide meaningful information to gain better understanding of CDSSs and to inform their successful adoption, implementation, usage and to justify the resource allocation. This research shows that CDSS evaluations investigated took a predominantly narrow view. It thus provides evidence for the need for a more systemic approach to evaluation.

# Contents

List of appendices	i
List of figures	i
List of tables	iv
List of abbreviations and acronyms	vi

<b>1 Introduction</b>	<b>1</b>
<b>1.1 The NHS and the challenges facing it</b>	<b>2</b>
1.1.1 Clinical Commissioning Groups	3
1.1.2 Primary Care Services	4
1.1.3 Acute NHS Trusts	4
<b>1.2 Challenges facing the NHS</b>	<b>6</b>
<b>1.3 Role of technology in meeting the demands placed upon the NHS</b>	<b>7</b>
1.3.1 Making effective use of technology within the NHS	9
1.3.2 Knowledge gap	10
1.3.3 Thesis aims and objectives	11
1.3.4 Research questions	12
<b>2 Literature Review</b>	<b>13</b>

<b>2.1 Introduction</b>	<b>13</b>
<b>2.2 Evaluation models and definitions</b>	<b>13</b>
<b>2.3 Approaches and methods of evaluation</b>	<b>17</b>
2.3.1 Theory based evaluations	17
2.3.2 Decision making evaluation models	18
2.3.3 Naturalistic evaluation approaches	21
2.3.4 Information Systems evaluation approaches	23
2.3.5 Health information systems evaluation approaches	25
<b>2.4 Evaluation of clinical decision support systems</b>	<b>30</b>
2.4.1 Types and taxonomy of CDSS	32
2.4.2 Benefits and uses of CDSS	35
2.4.3 Approaches to CDSS evaluation	38
2.4.4 Strategies for CDSS evaluation	42
2.4.5 CDSS evaluation methods	44
2.4.6 Pilot evaluation studies	45
2.4.7 Before and after evaluation studies	46
2.4.8 Controlled before and after evaluation studies	47
2.4.9 The Randomised Controlled Trial (RCT)	47
2.4.10 Systematic Reviews	49
2.4.11 Formative evaluations	65

---

2.4.12 Summative evaluations	65
2.4.13 Monitoring as a form of evaluation	65
2.4.14 Frameworks for CDSS evaluation	66
<b>2.5 Contextual and organisational issues related to evaluations</b>	<b>67</b>
2.5.1 Existing technology infrastructure	69
2.5.2 Organisational readiness for technology adoption	70
2.5.3 External technology context	71
<b>2.6 Purposes of CDSS evaluations</b>	<b>71</b>
<b>2.7 Benefits of CDSS evaluations</b>	<b>73</b>
<b>2.8 Barriers to CDSS evaluations</b>	<b>75</b>
<b>2.9 Trends in CDSS evaluation</b>	<b>81</b>
<b>2.10 Key findings of the literature review</b>	<b>87</b>
<b>3 Methodology</b>	<b>95</b>
<b>3.1 Introduction</b>	<b>95</b>
3.1.1 Research questions	96
<b>3.2 Research strategies considered</b>	<b>96</b>
<b>3.3 Justification for chosen research strategy</b>	<b>99</b>
<b>3.4 How the research was carried out</b>	<b>101</b>
3.4.1 Case selection	101
3.4.2 Data collection	103

3.4.3 Ethics approval	108
<b>3.5 Data analysis techniques</b>	<b>109</b>
3.5.1 Within case analysis	110
3.5.2 Cross case analysis	112
<b>3.6 Reflection on conduct of research</b>	<b>113</b>
<b>4 Background to study sites</b>	<b>115</b>
<b>4.1 The regulatory structure of health care in England</b>	<b>115</b>
4.1.1 Parliamentary regulation of healthcare	118
4.1.2 Independent regulatory bodies	119
<b>4.2 Professional regulation and collegiate societies</b>	<b>124</b>
<b>4.3 Introduction to the study Trust</b>	<b>127</b>
4.3.1 Selected clinical decision support systems	129
<b>5 T1 Risk Assessment Tool</b>	<b>130</b>
<b>5.1 Introduction</b>	<b>131</b>
5.1.1 Venous thromboembolism clinical interventions	131
<b>5.2 Overview of VTE Management in the NHS</b>	<b>132</b>
5.2.1 Guidelines for prevention and treatment of VTE	133
5.2.2 Patient safety initiatives for VTE prevention	136
5.2.3 Dr Foster Intelligence hospital ratings	138
5.2.4 Incentives for VTE management in the NHS	139

<b>5.3 Management of VTE in the study Trust</b>	<b>144</b>
5.3.1 Roles and responsibilities of key stakeholders	146
5.3.2 Paper based VTE risk assessment tool	147
5.3.3 The Technology	152
5.3.4 Implementation of T1 in the study Trust	154
<b>5.4 T1 evaluations</b>	<b>158</b>
5.4.1 Peer review of study Trust's VTE operational policies	162
5.4.2 Peer review of paper-based VTE rules and algorithms	162
5.4.3 Validation of T1 software rules	163
5.4.4 VTE tool usability and user acceptance testing	164
5.4.5 VTE compliance audits	164
5.4.6 Enoxaparin® and Doppler scanning cost-effectiveness audits	167
5.4.7 External CQUIN compliance audit	168
5.4.8 Root cause analysis of hospital acquired VTE	169
5.4.9 The Enoxaparin prophylaxis and treatment doses audit	172
<b>5.5 Discussion</b>	<b>178</b>
5.5.1 Purposes of evaluations <sup>1</sup>	178
5.5.2 Evaluation approaches and methods	179
5.5.3 Contextual and organisational issues related to evaluation	181
5.5.4 Barriers to evaluation	184

5.5.5 Overview of T1 evaluations and missed Opportunities	185
<b>6 T2 risk assessment tool</b>	<b>188</b>
<b>6.2 Introduction</b>	<b>189</b>
6.1.1 Prostate cancer interventions	190
<b>6.2 Prostate cancer: the UK clinical context</b>	<b>191</b>
<b>6.3 Management of stable prostate cancer in the study Trust</b>	<b>195</b>
6.3.1 The technology	198
6.3.2 How T2 works	199
6.3.3 Expected benefits of T2	201
<b>6.4 Implementation of T2 in the study Trust</b>	<b>205</b>
<b>6.5 T2 evaluations</b>	<b>208</b>
6.5.1 Validation of software rules	210
6.5.2 Clinical evaluation of the system	211
6.5.3 Pre-T2 adoption evaluations	214
6.5.4 T2 post market surveillance audit	216
6.5.5 T2 patient satisfaction survey	217
<b>6.6 Discussion</b>	<b>220</b>
6.6.1 Purposes of T2 evaluations	220
6.6.2 Approaches and methods used to evaluate T2	222



6.6.3 Contextual and organisational issues that affected T2 evaluations	226
6.6.4 Barriers to T2 evaluations	227
6.6.5 Benefits of T2 evaluations	229
6.6.6 Overview of T2 evaluations	230
<b>7 T3 track and trigger tool</b>	<b>232</b>
7.1 Introduction	233
7.2 Interventions for acutely ill medical patients	234
7.3 Management of acutely ill patients in the NHS	236
7.4 Management of acute medical patients in the study Trust	241
7.4.1 The technology	243
7.4.2 How T3 works	244
7.4.3 Expected T3 benefits	247
7.4.4 T3 implementation in the study Trust	248
7.5 T3 evaluations	260
7.5.1 Evaluations during implementation	263
7.5.2 Benefits realisation studies	265
7.5.3 IQNARC audits	267
7.5.4 Dr Foster mortality audits	269
7.5.5 Critical Care Outreach service evaluation	271

---

7.5.6 Root cause analyses for VTE	275
7.5.7 T3 user group network	277
7.5.8 Informal evaluations	279
<b>7.6 Discussion</b>	<b>283</b>
7.6.1 Purposes of T3 evaluation	283
7.6.2 Approaches and methods used to evaluate T3	285
7.6.3 Contextual and organisations issues	288
7.6.4 Barriers to evaluation	290
7.6.5 Benefits of T3 evaluations	292
7.6.6 Issues missed by T3 evaluations	293
<b>8 Cross Case Analysis</b>	<b>297</b>
8.1 Introduction	297
8.2 CDSS evaluations that were carried out	297
8.3 Key contextual factors affecting evaluations	303
8.3.1 Assessment of existing technology infrastructure	304
8.3.2 External pressures for CDSS evaluations	308
8.3.3 The attitudes of key stakeholders towards evaluations	310
8.4 Purposes of evaluations	314
8.5 Approaches and methods of evaluations used	318
8.6 Barriers to evaluations	322

8.6.1 Human barriers to evaluations	324
8.6.2 Methodological barriers to CDSS evaluations	327
8.6.3 Contextual and organisational barriers to CDSS evaluations	329
<b>8.7 Benefits of evaluations</b>	<b>331</b>
<b>8.8 Further factors of evaluations identified in the cases</b>	<b>332</b>
8.8.1 Funding for evaluations	333
8.8.2 Characteristics of the key stakeholders involved with evaluations	338
8.8.3 Criteria used to judge the success or failure of evaluations	341
<b>8.9 Limitations of the evaluations</b>	<b>343</b>
<b>8.10 Overview of CDSS evaluations that were carried out</b>	<b>346</b>
<b>9 Discussion and Conclusions</b>	<b>347</b>
9.1 Introduction	347
9.2 Research results	350
9.2.1 Contextual and organisational and social issues that affect CDSS evaluations	350
9.2.2 Purposes of CDSS evaluations	354
9.2.3 Approaches and methods of CDSS evaluations	361
9.2.4 Benefits of CDSS evaluations	369
9.2.5 Barriers to CDSS evaluations	374

---

<b>9.3 Research findings that support the literature review model</b>	<b>382</b>
9.3.1 Organisational, contextual and social issues that support the literature review model	382
9.3.2 Purposes of evaluations that support the literature review model	387
9.3.3 Approaches and methods for CDSS evaluations that support the literature review model	388
9.3.4 Benefits of evaluations that support the literature review model	392
9.3.5 Barriers to CDSS evaluations that support the literature review model	393
<b>9.4 Key research findings</b>	<b>398</b>
<b>9.5 Evaluation factors that extend the CDSS evaluation framework</b>	<b>405</b>
<b>9.6 Key contributions to CDSS evaluation literature</b>	<b>409</b>
9.6.1 Summary of research contributions to CDSS evaluation literature	413
<b>9.7 Potential limitations of this research</b>	<b>417</b>
<b>9.8 Conclusions and recommendations for further research</b>	<b>418</b>
<b>10 References</b>	<b>422</b>

## List of appendices

Appendix 1 – List of key study Trust documents and secondary materials	456
Appendix 2 – General indicative interview topics and questions	458
Appendix 3 – Guide to evaluation inputs, processes and outcomes	459
Appendix 4 - Participant Information Sheet	461
Appendix 5 – Consent Form	464
Appendix 6 – HREC Ethics Approval	465
Appendix 7 – GfREC Guidance and Study Trust R&D approval	467
Appendix 8 – Examples of coding practices and cross case analysis	470

## List of figures

Figure 1.1 Providers and commissioners of NHS services in England (adapted from NHS Confederation, 2015)	2
Figure 1.2 The NHS commissioning cycle (adapted from Adlington et al. (2013)	3
Figure 1.3 Barriers to innovation in the NHS (from Liddell, Ayling and Reid (2011))	10
Figure 2.1 Structure of CDSS (adapted from Berner, (2007))	33

Figure 2.2 Key factors for CDSS evaluation identified in the literature	89
Figure 4.1 Key functions of the Department of Health	116
Figure 5.1 Department of Health national VTE risk assessment template (The Department of Health, 2010)	136
Figure 5.2 Study Trust's electronic VTE risk assessment operational processes and work streams	145
Figure 5.3 Basic steps of a Root Cause Analysis (adapted from the NPSA (2014))	172
Figure 6.1 Timeline showing key prostate cancer initiatives and T2 adoption	193
Figure 6.2 Problems that were expected to be resolved by implementing and using T2	197
Figure 6.3 Respondents who participated in T2 telephone clinics (showing comparison between developing and adopting Trusts)	218
Figure 6.4 T2 exploratory survey results showing comparison between respondents from the developing and study Trusts	220
Figure 7.1 Problems that were expected to be solved by implementing and using T3	244
Figure 7.2 T3 core applications	245

Figure 9.1 Key factors of evaluation identified in the literature review	349
Figure 9.2 CDSS evaluation contexts as perceived by the key stakeholders	350
Figure 9.3 Factors affecting the purposes of CDSS evaluations and the key questions that need to be considered by the evaluators	360
Figure 9.4 Focus of evaluations at different stages of CDSS lifecycle and the key questions which need to be considered by the evaluators	368
Figure 9.5 Hierarchy of benefits of CDSS evaluations identified in the cases and the key questions that need to be considered by evaluators	373
Figure 9.6 Synthesised model of the key factors that affect CDSS evaluations	411

**List of tables**

Table 2.1 Stufflebeam’s CIPP model	19
Table 2.2 Characteristics of three major types of evaluation study designs [Adapted from Wyatt and Wyatt (2003)]	44
Table 2.3 Summary of systematic reviews of CDSS evaluations (adapted from Cochrane database of systematic reviews)	51
Table 2.4 Literature supporting the CDSS evaluation framework	92
Table 3.1 Relevant situations for different research methods [Adapted from Yin (2009)]	97
Table 5.1 VTE risk assessment data collection for NHS acute providers (the Department of Health)	141
Table 5.2 Evaluations undertaken reported by participants in chronological order	160
Table 5.3 NICE VTE Prevention Quality Standard (adapted from NICE)	174
Table 5.4 Opportunities that were missed by the evaluations	187
Table 6.1 T2 evaluations in chronological order	209
Table 6.2 Opportunities that were missed by T2 evaluations	231
Table 7.1 Duties and responsibilities of key actors involved with T3	250



Table 7.2 Evaluations undertaken in the study Trust (in  
chronological order) 261

Table 7.3 Issues missed by T3 evaluations 295

Table 8.1 CDSS evaluations carried out before, during  
and after implementation 299

Table 8.2 Factors to consider when assessing existing technological  
infrastructure to support CDSS evaluations 305

Table 8.3 Purposes of CDSS evaluations at different phases of their  
adoption lifecycle 316

Table 8.4 Evaluation approaches and methods used 319

Table 8.5 Barriers to CDSS evaluations 323

Table 8.6 Resources allocated for CDSS project teams 335

Table 8.7 CDSS evaluators and target audiences 339

Table 9.1 Key evaluation questions that could guide which CDSS  
evaluation methods to use 363

Table 9.2 Barriers and enablers CDSS evaluations 379

Table 9.3 Summary of key research findings 399

## **List of abbreviations and acronyms**

AHRQ	Agency for Healthcare Research and Quality
BBC	British Broadcasting Corporation
BSH	British Society for Haematology
CCG	Clinical Commissioning Groups
CDSS	Clinical Decision Support Systems
CIPP	Context, Input, Process and Product
CQUIN	Commissioning for Quality and Innovation
EPRS	Electronic Patient Record System
FDA	Food and Drug Administration
GEP-HI	Good Evaluation Practice in Health Informatics
GMC	General Medical Council
GP	General Practitioner
HIS-EVAL	Health Informatics Systems Evaluation
HREC	Human Research Ethics Committee
HSCIC	Health and Social Care Information Centre
HSMR	Hospital Standardised Mortality Ratios
HTA	Health Technology Assessment
ICF	Information and Computer Technology
ICF	Informed Consent Form
ICNARC	Intensive Care National Audit and Research Centre

ICT	Information and Communication Technology
ICU	Intensive Care Unit
IMIA	International Medical Informatics Association
INAHTA	International Network of Agencies for Health Technology Assessment
MEWS	Modified Early Warning Score
MHRA	Medicines and Healthcare Products Regulatory Agency
NAIP	National Infract Angioplasty Project
NCAA	National Cardiac Arrest Audit
NEWS	National Early Warning Score
NHS	National Health Service
NICE	National Institute for Clinical Excellence
NMC	Nursing and Midwifery
NPfIT	National Programme for Information Technology
NPSA	National Patient Safety Agency
OCS	Order Communication system
PDA	Personal Digital Assistant
PIS	Participant Information Sheet
RCT	Randomised Controlled Trial
TDA	Trust Development Authority

UK	United Kingdom
USA	United States of America
VTE	Venous Thromboembolism

# Chapter 1 Introduction

This thesis looks at the evaluations of three computerised clinical decision support systems that were undertaken in a UK National Health Service (NHS) Trust. A clinical decision support system (CDSS) is software that uses an inbuilt knowledge base to help physicians, healthcare staff, patients and carers to make informed clinical decisions at the point of care based on person-specific information. CDSSs aim to enhance healthcare delivery by using tools such as alerts, suggestions, reminders, and prompts to help users make better informed decisions. The purpose of the research was to systematically assess the evaluations of CDSSs that were implemented by a large NHS Trust in order gain an understanding of how these systems were evaluated and the wider implications of these evaluations.

This chapter provides an overview of the NHS and its drive to adopt computerised healthcare information systems to support clinical processes and improve clinical decision-making. Chapter 2 reviews the literature that is relevant to this research and Chapter 3 presents the methodology that was used to undertake it. Chapter 4 looks at the background to the settings in which the research was undertaken and Chapters 5, 6 and 7 present case studies of the evaluations of three different CDSSs, and Chapter 8 provides a cross case comparison. Chapter 9 presents the findings of the research and looks at them in the context of the most relevant factors of the literature review before making suggestions for future research.

In this chapter, Section 1.1 describes how the NHS was established and how it is organised. It also highlights some of the key challenges facing the NHS. The role that computerised health information systems were expected to play in supporting the NHS is discussed in

Section 1.2. The pressures faced by NHS organisations to effectively use these systems are discussed in Section 1.3.

## 1.1 The NHS and the challenges facing it

The UK NHS was established in 1948 following social reforms that sought to deliver universal healthcare for free at the point of delivery. Since its founding, the NHS has been a battleground for all political parties. It has faced many reforms relating to its structure and funding through various modernisation and performance improvement initiatives. The NHS is divided into primary and secondary care services. Figure 1.1 shows the different providers and commissioners of NHS services in England. Their roles will be discussed in turn.

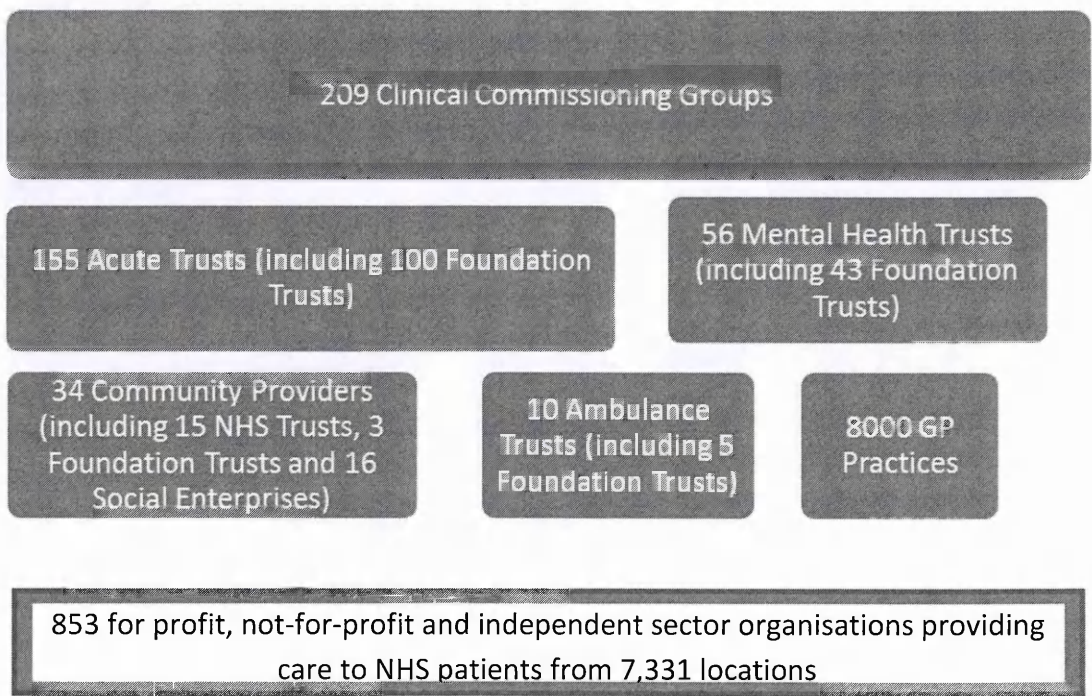


Figure 1.1 Providers and commissioners of NHS services in England (adapted from NHS Confederation, 2015)

1.1.1 Clinical Commissioning Groups

The traditional NHS system had several NHS Health Trusts that managed hospital care, community care and mental health services in England. The Department of Health, Strategic Health Authorities and Primary Care Trusts had responsibilities for all planning and delivery of NHS services. These statutory responsibilities were transferred to NHS England in April 2013. The NHS Trust Development Authority (TDA) now provides support, oversight and governance for all NHS Trusts in England. Clinical Commissioning Groups (CCGs) are groups of General Practices who have a responsibility for the planning and designing of healthcare services for local populations. There are 209 CCGs in England. Their commissioning role involves the allocation for and purchasing of health and care services including: urgent and emergency care; planned hospital procedures; mental health and learning disability; rehabilitation care and community services. NHS England has overall responsibility to ensure that CCGs adequately commission healthcare services for their local populations and also that they meet their financial obligations. Figure 1.2 shows the key elements of the NHS commission cycle.

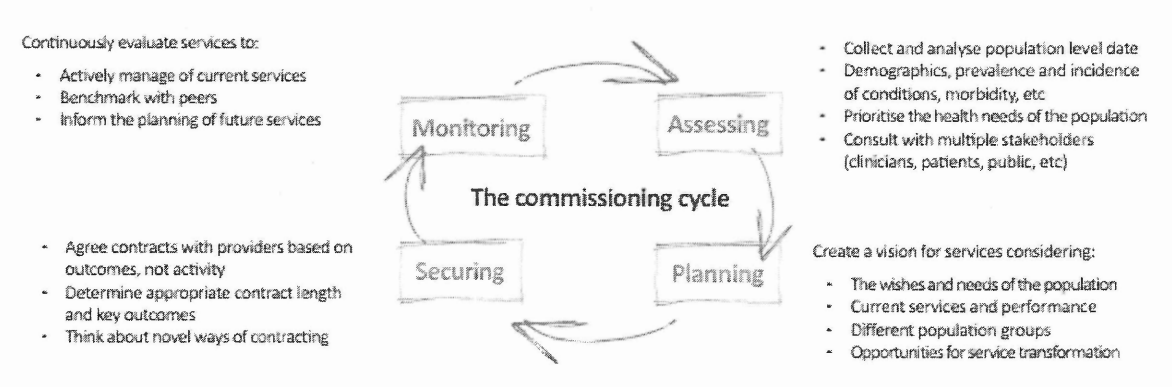


Figure 1.2 The NHS commissioning cycle (adapted from Adlington et al. (2013)

### **1.1.2 Primary Care Services**

Primary Care services provide the first point of contact for most people seeking access to healthcare services in England. From 1948 to 2004, every GP principal was responsible for delivering care to registered patients within their practice and following up those with long-term conditions. Since 2004, patients were required to register with a General Practice rather than individual GPs. Additionally, patients can also access care through NHS call centres (such as the NHS 111 telephone services), walk-in centres, GP-led out of hours clinics, district and community nursing services, private sector organisations that are contracted by the NHS and voluntary sector organisations among others. There are around 8000 GP Practices in England and 853 third sector organisations (for profit and not for profit) that provide healthcare services to patients from 7331 locations across England.

### **1.1.3 Acute NHS Trusts**

NHS secondary care services include acute district hospitals, Acute Teaching NHS Trusts and NHS Foundation Trusts. There are 155 acute NHS Trusts in England, including 100 NHS Foundation Trusts. Their services are primarily commissioned by the CCGs within their local areas. Acute NHS Trusts are responsible for managing hospitals and ensuring that they provide high quality care and financial management. They are also responsible for employing multi-disciplinary clinical and non-clinical workforce, including doctors, nurses, physiotherapists, allied healthcare professionals, administrative staff, cleaners, managers and information technology technicians amongst many others. Acute Teaching NHS Trusts are linked with local universities to provide training for clinical professionals such as doctors, nurses and scientific staff among others. Some Acute NHS Trusts provide specialist services and may act as regional or national centres for certain conditions within their hospitals. They may also provide community based care in partnership with community



clinics and health centres. NHS Foundation Trusts were introduced in April 2004. Unlike Acute NHS Trusts, they are independent legal entities that are accountable to the populations they serve under unique governance arrangements rather than direct accountability to the government. They also have financial freedoms that allow them to raise capital from both private and public sectors and to set their own performance criteria. NHS Foundation Trusts are required to work closely with their boards of governors, who include members of staff, patients and local partners. All NHS Foundation Trusts are overseen by Monitor, the new NHS regulatory body.

There are 56 Mental Health NHS Trusts, including 43 Foundation Trusts. Mental Health NHS Trusts provide health and social care services for people with mental health problems. These services can be provided through mental health hospitals, GPs, specialist care services and other primary care providers. There are 34 community care providers that provide non-acute health and social care services in England. They include 15 NHS Trusts, 3 Foundation Trusts and 16 social enterprises. There are 10 Ambulance Trusts in England, including 5 Foundation Trusts, which provide access to emergency healthcare from primary to secondary care. The NHS Trust Development Authority (TDA) helps Acute NHS Trusts through their transition to Foundation Trust status. The TDA took over the responsibility for the governance, quality and performance management of Acute NHS Trusts following the abolition of Strategic Health Authorities in 2013. NHS England and CCGs also get support and advice from Clinical Senates and Strategic Clinical Networks. Clinical Senates are advisory groups that are made up of clinical leaders and experts from various healthcare, social care and public health settings and sometimes patients. There are 12 Strategic Clinical Networks in England. They provide expert advice on particular diseases and help to improve clinical pathways that can be adapted into clinical practice to help

reduce variations in healthcare services. They also provide guidance on innovations that need to be adopted, advice to clinical leaders to support wider decision-making and strategic planning. The NHS also has special health authorities which provide health services across England. They include the National Blood and Transplant Authority, NHS Business Services Authority and the NHS Litigation Authority. Although they work independently, these authorities may be subject to ministerial direction where necessary.

## **1.2 Challenges facing the NHS**

Over the past two decades, there has been a shift in disease patterns from acute to chronic conditions, with conditions such as diabetes, hypertension and asthma requiring systematic monitoring in dispersed community settings. There has also been a significant increase in life expectancy in the UK. This has led to challenges relating to an aging population, such as multiple chronic illnesses (comorbidity) and the related prescription of multiple medications (polypharmacy), which come with additional risks of drug interactions and allergies (Greenhalgh, 2007). This group of patients has become particularly burdensome for the NHS to manage because they require constant trans-disciplinary follow up between both primary and secondary services. There have been calls to enhance the integration of primary and secondary care organisations to allow collaboration on innovative service delivery models and more effective communication to reduce duplication and to ensure continuity of care. This has led to the development of integrated services, where geographically dispersed multi-disciplinary teams work together to deliver services. Concerns about the quality of healthcare and patient safety in the 1990s and the early 2000s led to the new public management era, which emphasised a growing culture of accountability, surveillance, regulation and performance management in NHS

organisations. This led to increased scrutiny of all clinical processes. NHS Trusts were required to undertake and publish audits relating to various aspects of patient care. Regulatory bodies such as the National Institute of Health and Clinical Excellence and the National Patient Safety Agency and professional bodies such as the Royal Medical Colleges have led the efforts to enforce this culture at organisational and individual levels

## **1.3 Role of technology in meeting the demands placed upon the NHS**

Following the 1997 general election, the Labour government introduced reforms that were aimed at modernising public services through service integration, increased efficiencies, promoting accountability and transparency and delivering user-focused services (Darling, 1998; Newman, 2006). They sought to achieve these extensive NHS modernisation goals through large-scale IT systems and empowering service-users to make choices about their health. The modernisation drive was given a boost by Sir Derek Wanless's report, *Securing Our Future Health* (Wanless, 2002), which recommended that the NHS should substantially increase investment in integrated large-scale IT systems to support the required improvements in the quality of care and efficiency. One such large-scale project was the National Programme for IT (NPfIT), which benefitted from unprecedented economic growth from the late 1990s and early 2000s which helped to fund public services. However, like many developed economies after the financial crisis in 2008, the UK entered a deep recessionary period, which prompted the government to rethink NHS and wider public sector funding. In 2009, the Department of Health subsequently issued a five year plan which required NHS Trusts to save nearly £20 billion by 2014 while at the same time improving quality of care and efficiency. The Conservative-Liberal Democrat coalition

government continued these austerity measures following their election in 2010, leading to sustained pressure on the NHS to implement more cost-saving measures without compromising patient care.

There is widespread agreement that demand for NHS care will continue to grow due to factors such as the increasing proportion of elderly people in the population, medical advances and the rising expectations of the public for medical investigation and treatment. However, this increase in demand will not be matched by increases in resources. The combined consequences of greater demand and economic pressures will mean that successful innovation and continuous improvement will become even more essential if the health systems are to stand a chance of meeting the demands placed upon them. The Carruthers review and subsequent report (Liddell et al., 2011) focused on the need to encourage innovation of services, especially in the context of the push for improved quality and productivity. Many of the service changes the report identified as necessary rely on the development and implementation of innovative technologies. In order to produce these technologies it is necessary for the health technology innovation process to be able to satisfy previously ignored or unarticulated clinical needs rapidly and precisely.

Liddell, Ayling and Reid's report entitled *Innovation Health and Wealth, Accelerating Adoption and Diffusion in the NHS* (Liddell et al., 2011: 7) points out that:

The purpose of the NHS, and everyone working in it, is to promote health and wellbeing, and to provide high quality healthcare, free at the point of delivery to everyone who needs it. ... Innovation has a vital role to play in fulfilling this purpose by improving the quality of care for patients, releasing savings through productivity, and enabling the NHS to make its contribution as a major investor and wealth creator in the UK.

However, despite major technological advances over the past three decades and their potential and demonstrated benefits, very few computerised healthcare information technologies are widely adopted and used in NHS clinical settings. Cooksey (2006) noted that healthcare organisations often struggled to adopt innovative technologies. Similarly, Lord Darzi, in his 'once in a generation review' of the health service in England noted:

The NHS does not always make best use of innovation. ... Despite some excellent work taking place locally, there remains some reluctance within the NHS to adopt new products and processes

Darzi (2007: 40)

The NHS technology landscape has often shifted in response to the technological advances, and is driven by the government through its arm's length bodies and regulatory authorities, as well as developers and commercial vendors. Increasing media and public scrutiny of NHS services has led to pressures on NHS organisations to deliver high quality services and adopt new ways of working within an increasingly cost-constrained environment.

### **1.3.1 Making effective use of technology within the NHS**

Technology adoption in health services is often a complex process, especially in trans-disciplinary environments such as hospitals. These complexities may be related to the technologies themselves or specific problems in adopting such technologies. Traditional technologies such as pharmaceutical products have been reported to be much easier to adopt due to their self-containment and limited organisational effects (Department of Health, 2008; Eaton, 2008). In contrast, healthcare technologies such as surgical devices

and clinical decision support software have been found to be more complex to adopt because the intended users often need to gain new skills and knowledge about them and must also be willing to change their habits and practice. The wider organisational effects of such technologies may affect the existing professional hierarchies, result in shifts in responsibilities and bring about radical service redesign. Various NICE guidelines have recommended that NHS Trusts should implement clinical decision support systems to help frontline staff to make better informed decisions at the point of care. However, there are a range of barriers that affect the adoption of innovations in healthcare. Figure 1.3 shows some of the barriers to innovations that were identified in a Department of Health report.

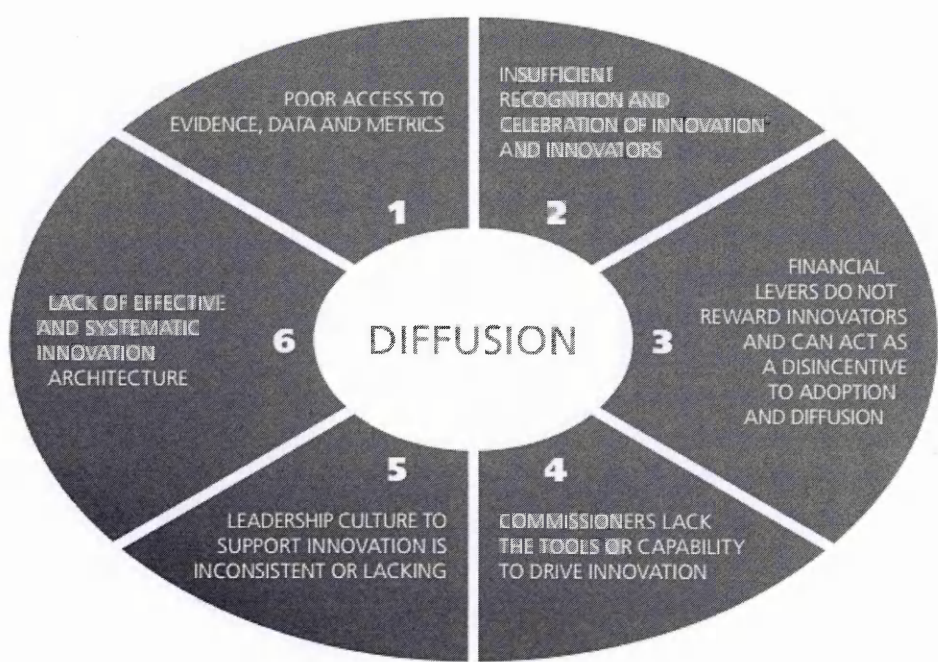


Figure 1.3 Barriers to innovation in the NHS (from Liddell, Ayling and Reid (2011))

### **1.3.2 Knowledge gap**

Since the 1960s, there have been repeated calls for a “gold standard” approach to evaluate the effectiveness of healthcare interventions, including clinical decision support systems. A response to these calls has been the so called “hierarchy of evidence” which is based on randomised control trials and other experimental designs. Some have argued, however, that experimental methods do not provide sufficient information regarding why some interventions work and others do not, and have called instead for an “eclectic” range of evaluation methods. It is fair to say that half a century later, disagreements still remain regarding what can and cannot be evaluated, the context of evaluation, the role of evaluation, the focus and methodologies for evaluation, and that there are still methodological, technological, organisational and human barriers to effective CDSS evaluations.

### **1.3.3 Thesis Aims and objectives**

The aim of this research project was to explore how CDSS evaluations are carried out in selected NHS settings with a view to adding new knowledge to the existing CDSS evaluation literature.

The objectives were to:

- 1) Critically appraise CDSS evaluation literature
- 2) Fully identify the gap in knowledge and develop appropriate research questions
- 3) Develop robust research methodologies and carry out the research
- 4) Systematically analyse data collected

- 5) Discuss the research findings and draw firm conclusions and recommendations for research and practice

### **1.3.4 Research questions**

In order to make this contribution to the CDSS evaluation literature, the following research questions were framed:

1. What are the key factors that affect CDSS evaluations in a typical NHS hospital setting?
2. How do these factors relate to the model of evaluation that was developed from the CDSS literature review?
3. To what extent do evaluations affect decisions to adopt CDSSs in healthcare settings and which evaluation methods are most likely to inform CDSS adoption decisions and why?

The next chapter will look the literature relating to the evaluation clinical decision support systems.



# **Chapter 2 Literature Review**

## **2.1 Introduction**

This chapter examines the literature that is relevant to research looking at the evaluation of clinical decision support systems and thus provides information about the context for the research as well as helping to identify the research opportunity more clearly and support the development of a research strategy. In Section 2.2, broad definitions of evaluation are given, noting their influences across various disciplines. Section 2.3 gives a wide overview of the evaluation landscape across various disciplines, highlighting key approaches such theory-based, decision making and naturalistic evaluation models, and information systems and health information systems evaluation approaches. In Section 2.4, clinical decision support systems are defined, noting their types and taxonomy, range of uses and the approaches and methods that are commonly used to evaluate them. Section 2.5 gives an overview of the key findings from the literature review and concludes with the research questions.

## **2.2 Evaluation models and definitions**

There are many models and definitions of evaluation. In this section, various definitions of evaluation are discussed in the context of established education-based evaluation models such as theory based evaluation, programme evaluation, decision making evaluation and evaluation approaches that are commonly used for healthcare interventions and information technology. Models used to evaluate education programmes in the 1960s and 1970s have shaped the wider field of evaluation over the last four decades, including the evaluation of healthcare interventions. Stufflebeam and Shinkfield (2007) noted that

because of the many evaluation approaches and activities, the definition of evaluation has changed over the years. Earlier definitions of evaluation primarily focused on the achievement of set objectives as defined by Tyler (1942). Tyler's study of schools in the 1930s led to the development of a model that sought to determine whether a program achieved its goals as the primary role of evaluation. However, Stufflebeam and Shinkfield (2007) rejected Tyler's definition, arguing that it steered evaluation activity towards the attainment of outcomes. They noted that some program objectives may be flawed and not worthy of achievement. They also argued that a broader evaluation approach looking at a program's goals, structure and process was required, particularly in cases where program improvement or wider adoption/adaptation by other service providers was paramount.

Evaluation models that were developed to provide accountability in large healthcare, education and welfare reform programs in the USA in the 1970s looked beyond Tyler's approach (Stufflebeam, 1973b; Stufflebeam, 1973a; Scriven, 1972). This led to a shift in the definition of evaluation towards professional judgement, particularly the collection and analysis of information that can be used for decision making (Stufflebeam and Shinkfield, 2007). Suchman (1967) noted that although evaluation of a program's goals remained important, the intervening processes that led to the observed outcomes were more essential. Schuman further argued that evaluations should test hypotheses regarding a program's activities, objectives and processes. Similarly, Scriven (1972) argued that other outcomes were equally as important as the programme's stated goals, noting that evaluations should look at both intended and unintended outcomes. He advised evaluators to avoid the rhetoric about programs such as brochures, proposals and other professional materials and instead only focus on the actual outcomes. Scriven (1972) also popularised formative and summative evaluations in the 1970s to distinguish between the two roles

that evaluators can play during a programme's lifecycle. Formative evaluations focus on the assessment of the program during its development while summative evaluations assess the outcomes of a completed program (Scriven, 1972). However, Stufflebeam and Shinkfield (2007) argued that distinguishing these two roles was not as clearly demarcated in practice as suggested by Scriven, although this distinction continues to be widely drawn on. They argued that evaluation should instead focus on collecting and analysing quality information for decision makers (Stufflebeam and Shinkfield, 2007).

Other definitions of evaluations are based on the methodology employed, for example experimental designs which look at the comparative outcomes of different programs or interventions, services or products (Stufflebeam and Shinkfield, 2007). Some argue that establishing causality is the primary purpose of evaluation and that it can only be achieved through experimental approaches (Cochrane, 1972; Liu and Wyatt, 2011b). Cochrane (1972), advocated for wider use of randomised controlled trials as the "gold standard" for evaluating health and social care interventions to assess their effectiveness and efficiency. Cochrane's work throughout the 1970s and 1980s resulted in the establishment of the Cochrane library database of systematic reviews and the elevation of randomised controlled trials to the top of the "hierarchy of evidence" in evaluating outcomes of healthcare interventions (Evans, 2003). Using this model, assessing the quality or effectiveness of healthcare interventions or health services is measured through predefined outcomes, which can be at an individual level, for example quality of life improvement or system level such as cost-effectiveness or disease burden (Jefford et al., 2003). However, Donabedian's (2005) paradigm for quality and outcome assessment recognised that some outcomes such as patient attitudes, satisfaction and social restoration may not be easily defined and measured. Stufflebeam and Shinkfield (2007: 8)

also noted that controlled experimental designs were not feasible in most cases, and that they could be counterproductive or fail to address key issues about “needs, goals, plans, processes, side effects” among others. They argued that evaluators should consider a wide range of available methods to “reach defensible judgments of programs” rather than equate with any one particular method.

Another widely used definition of evaluation is that of the United States Joint Committee on Standards for Educational Evaluation (1994: 3) which stated that “evaluation is the systematic assessment of the worth or merit of an object”. According to Stufflebeam and Shinkfield (2007), this definition entails value judgments and thus requires a set of defensible guiding principles. Accordingly, these guiding standards may also be used to assess the merit of a program or intervention. However, evaluations such as formative and descriptive studies do not necessarily assess the merit or worth of an object. Ammenworth et al. (2004: 2) defined evaluation as “the decisive assessment of defined objects, based on a set of criteria, to solve a given problem”. This definition is focused on resolving a given problem through a systematic assessment of defined objects that contribute to the problem. Similar to Stufflebeam and Shinkfield’s (2007) approach, Ammenworth et al.’s (2004) definition focused on the collection and assessment of information related to the problem, then assessing the worth rather than defining evaluation according to a programme’s worth.

This section has shown that there is no widely accepted definition of evaluation. Some definitions focus on the stated objectives of a program or intervention, while others depend on the context and methods used. Stufflebeam and Shinkfield (2007) suggested that evaluators should adopt an eclectic approach that considers the objectives, context of evaluation and various methods to develop and guide programs towards successful

outcomes. This entails an important role for evaluation that goes beyond just providing feedback to decision makers at the end of the program.

## **2.3 Approaches and methods of evaluation**

This section looks at the approaches and methods of evaluation identified in the literature from the disciplines of education, information systems, health and social care and clinical decision support systems.

### **2.3.1 Theory based evaluations**

Weiss (1972) advocated for evaluation models to be developed and tested to assess the ultimate objective of a program through an assessment of a chain of events. This early work by Weiss led to the development of program theory evaluation. According to Rogers et al. (2000a), program theory evaluation consists of two basic elements. Firstly, developing an explicit theory or model of how the programme causes the intended observed outcomes. Secondly, performing an evaluation guided by the model. According to Coryn et al. (2011), evaluation theories describe and prescribe how evaluations should be conducted, including such issues as purposes, users and uses of evaluations, who performs the evaluation process and to what extent, choices of strategies and methods, as well as the roles and responsibilities of evaluators amongst other things (Fournier, 1995; Smith, 1993). Chen's (1990) seminal book brought the concept of theory driven evaluations back to prominence, leading to further conceptual, methodological and theoretical contributions (Chen and Rossi, 1992; Rogers et al., 2000a; Rogers et al., 2000b; Weiss, 1997; Weiss, 1995; Weiss, 1998; Weiss, 2004). Theory driven evaluation has also been referred to as programme theory evaluation, theory based evaluation, theory guided evaluation, theory of action, theory of change, program logic, logical frameworks, outcomes hierarchies, and latterly

realist or realistic evaluation (Mark et al., 1998; Pawson and Tilley, 1997; Rogers et al., 2000a; Rogers et al., 2000b) and program theory evaluation science (Donaldson, 2007).

Instead of asking “what works” or “does this programme work”, realist evaluation seeks to ask “what works for whom in what circumstances and in what respects, and how?” (Pawson and Tilley, 1997). Pawson and Tilley argued that programs are theories that are embedded into the social systems where they are delivered. They also noted that programs were active in the sense that they are produced and require the engagement of individuals in an open system that is often affected by various external events (Pawson and Tilley, 1997). Their “context-mechanism-outcome” (CMO) configuration breaks down systems into their key components and processes to gain an understanding of social interactions which enable the evaluator to test and refine theories. This approach was used by Greenhalgh et al. (2009) in an organisational case study evaluating large scale service transformation in healthcare. Through realist evaluation, they drew useful lessons about how specific preconditions make specific outcomes likely, although it did not necessarily produce predictive guidance for the evaluators. Realist evaluation considers the complexity of programs or interventions to determine “what works for whom” rather than just looking at the stated objectives or cause and effect as advocated by proponents of experimental methods. Theory-based evaluations marked a significant shift from objectives-based and outcomes-based evaluations. Instead, they focused on key contextual and other contributory factors that resulted in the observed outcomes.

### **2.3.2 Decision making evaluation models**

Also in the 1970s, some researchers developed decision-making evaluation models which focused on how evaluation results were used. Stufflebeam (1973a) viewed evaluation as a

process that provided decision makers with meaningful and useful information regarding decision alternatives. His “context, input, process, and product” (CIPP) model described four types of evaluation activities. These evaluation activities are shown in Table 2.1. The CIPP model thus systematically assesses what needs to be done, how it should be done, whether it is being done and whether the program is successful. Stufflebeam (2007) also stated that Scriven’s formative evaluation related primarily to decision making, while summative evaluation was more focused on accountability.

Evaluation type	Evaluation activities
Context evaluation	Assessment of the problems, needs and opportunities available in the program’s setting
Input evaluation	Assessment of competing strategies and the work plans and budgets
Process evaluation	Monitors, documents and assesses a program’s activities
Product evaluation	Examines the program’s impact on the target audience, the quality and importance of outcomes and the program’s sustainability and transferability

Table 2.1 Stufflebeam’s CIPP model

Patton's (1978) utilisation focused evaluation model addressed the concern that decision makers often ignored evaluation findings by developing a general approach to evaluation based on two essential requirements. First, Patton argued that decision makers and audiences of evaluation reports should clearly be identified. Second, he recommended that evaluators should work with decision makers to decide about all aspects of evaluation, such as the evaluation questions, research design, data analysis, interpretation and dissemination of results. Cronbach (1980) also contributed to the discourse of decision-making evaluation. He stressed that decision-making was political and mostly involved various actors rather than lone decision-makers. Cronbach suggested that the evaluator should take the role of a teacher who educates and advises the client throughout the evaluation process. During this process, the evaluator should constantly give feedback to the client rather than just wait for the final report. Unlike Stufflebeam and Patton, Cronbach did not think it was the evaluator's role to determine the program's worthiness nor recommend appropriate courses of action. However, in practice the different roles of evaluators and decision makers may not be clearly distinguishable.

Another widely used education-based evaluation approach is the four level model developed by Kirkpatrick (1994). Although it primarily focuses on the corporate human resource development, it has been adapted to evaluate teacher professional development programs (Guskey, 2000) and the assessment of health information management courses and programmes (Rouse, 2011). This model is comprised of reaction, learning, behaviour, and results. Reaction refers to the measurement of participants' satisfaction with the program, usually through a survey. Learning relates to the measurement of the extent to which participants' attitudes change, their knowledge or skills improve through attendance of the program using course exams, tests or surveys. Behaviour relates to the measurement



of the extent to which participants' behaviour changes as a consequence of attending the course. This evaluation looks at the extent to which lasting change has occurred as a result of attending the course. The extent can be measured through improvements in productivity, management or quality. Kirkpatrick recommended the use of control group comparisons to assess a program's effectiveness at the behaviour and results levels. This model was influenced by scientific approaches such as randomised controlled trials, which maintain that effectiveness is reliably measured by eliminating bias in the evaluation process by controlling the environment (Cochrane, 1972). Decision making evaluation models shifted the emphasis of evaluation from outcomes and objectives of programs to other contributory factors in the lifecycle of a program or intervention. Such factors provide essential information regarding how a program or intervention works and how it results in the observed outcomes.

### **2.3.3 Naturalistic evaluation approaches**

Naturalistic evaluation approaches take the view that it is better to evaluate in the context of the phenomenon rather than in artificial conditions (Stake, 1975). They also allow the evaluator freedom to select appropriate methods to collect, analyse and interpret data. For example, Stake's (1975) responsive evaluation approach sought to address the concern that existing approaches did not sufficiently look at the needs of the evaluation client. Stake advocated that evaluators should learn from the anthropology and journalism traditions and use various qualitative methods to observe, collect data and report on program activities and satisfy the information needs of the target audience. He also recommended that evaluators should present the different value judgments. Eisner's (1979) connoisseurship model drew from the field of art criticism to support the evaluator's judgment of the quality of an educational program. This model is comprised of educational

connoisseurship and educational criticism. The former involves application of background experience to aid the evaluator in appreciating the merits of an educational program. The latter is dependent on the evaluator's ability to verbalise a program's features to make it easily understandable for those who lack the fine appreciation of connoisseurs. Guba and Lincoln's (1989) fourth generation evaluation approach highlighted the inherent problems with previous generations of evaluation approaches. They noted the difficulties in negotiating the political and ethical dilemmas, gaps, imperfections and often inconclusive deductions. They argued that the failure and lack of utilisation of evaluation results emanated from the unquestioned reliance on the positivist paradigm of research (Guba and Lincoln, 1989). Fourth generation evaluators seek to build an intricate understanding of the program they are studying through human, political, social, cultural and contextual issues that are involved. The focus of fourth generation evaluation is to build a rich description of the program and to communicate the appraisal of the program to its stakeholders. Data collection methods used in this approach include field notes, key informant interviews, case histories and surveys. Naturalistic evaluation approaches emphasise the significance of the natural evaluation context rather than creating an artificial environment for evaluation. They advocate for methodological pluralism to ensure that key issues such as contextual, human, political and cultural issues are taken into account in any evaluation activity.

### 2.3.4 Information Systems evaluation approaches

Mumford (1983) highlighted the social and technical nature of information systems and recommended the development of systems that would enhance job satisfaction. She argued that human factors were often not taken into account when introducing computer systems into organisations, resulting in failure to achieve the desired outcomes. Klecun and Cornford (2005) noted that the traditional focus of information systems' evaluation had been on immediate usability needs, their technical aspects, performance, reliability, robustness and security and cost-benefit assessments. They argued that evaluation of information systems remained "controversial, biased, incomplete and superficial" (Klecun and Cornford, 2005). Walsham (2006) argued that the increasingly pervasive and complex nature of information systems and their use as tools for organisational change entailed new demands from evaluation activities. Klecun and Cornford (2005) also noted that the complexity of new systems introduced new sociotechnical aspects that were difficult to evaluate. Land (2000) recommended that information systems evaluations should include wider perspectives at technological, economic, organisational and individual levels. Land suggested the use of a modified version of Kaplan and Norton's (1995) balanced scorecard to cover a range of issues such as customer, financial, internal efficiency, learning and growth and employees' perspectives. Martinsons et al. (1999) also looked at how the balanced scorecard could be adapted to the management of business functions, organisational units and individual information systems projects. Other researchers looked at aspects such as alignment of new technologies with strategic and business goals, understanding of formal and informal work structures, diversity and competing interests of stakeholder groups and their information needs (Symons, 1991; Smithson and Hirschheim, 1998). Cornford and Klecun-Dabrowska (2001) looked at ethical issues and the

effects of new systems on legitimate interests of stakeholders. Others found evidence of evaluation being used for legitimising vested interests (Friedman and Wyatt, 2006; Rigby, 2001).

The concept of benefits realisation management, or simply benefits management or benefits realisation in IT projects relates to how resources are utilised to make desirable improvements (Sapountzis et al., 2008). It also aims to align project outcomes with business strategies, and actively manage the benefits throughout the life cycle of the investment (Ward and Daniel, 2006). This concept was introduced in the late 1980s (Farbey et al., 1993) and has become an increasingly important factor in planning for how benefits from IT projects are realised and measured (Glynne, 2007). The increasing use and complexity of IT systems (Ashurst and Doherty, 2003; Ward and Elvin, 1999; Bradley, 2010) and the need to ensure that benefits from investments are realised and that decision makers are held to account have also contributed to the popularity of this method (Sapountzis et al., 2008). In the UK, the McCartney report, *Successful IT: Modernising Government in Action* and the CSSA report *Getting it Right for Government* (Wheeler-Carmichael, 2000) stated that programs can only be regarded as successful if their intended benefits are realised. Reiss (2006) also noted that most unsuccessful programs were characteristically vague regarding their expected benefits. Over the past decade, benefits realisation has gained popularity in healthcare settings because of the challenges in ascertaining whether benefits have been achieved due to complex interplay between multi-stakeholder teams and various IT systems. The benefits realisation model requires expected benefits to be outlined from the outset and a project is only considered successful if the intended benefits have been realised. A key element of this method is the need to consider the essential roles that each

stakeholder plays and manage all stakeholder groups and where necessary, assign benefits to different stakeholder groups.

Another widely used technology evaluation approach is health technology assessment (HTA). HTA is defined as a “tool to review technologies and provide evidence of the value these technologies can deliver to patients and their families, health system stakeholders, and to society more broadly” (INAHTA, 2015). According to Drummond et al. (2008), HTA seeks to inform decision makers regarding the benefits, risks and costs of new and existing technologies. This method has historically been used to produce information that can be used by a wide range of decision makers. However, the focus of HTAs has shifted to informing particular resource allocation decisions regarding healthcare technologies, evaluating costs and benefits as well as paying attention to the decisions influenced by the HTA (Drummond et al., 2008). There is increasing awareness of the importance of human, organisational and social factors that may affect the introduction of new technology (Kaplan, 2001a; Kaplan, 2001b). Kaplan (2001b) argued that there should be a shift in attitudes around health information technology evaluation from the traditional focus on technical and economic issues to sociotechnical issues. She noted that no single method could adequately provide decision makers with all the information that was required to make informed decisions about health information systems.

### **2.3.5 Health information systems evaluation approaches**

Haux et al. (2004) defined health information systems as systems that process data and knowledge in healthcare environments. According to Winter et al. (2001), health information systems include the overall information processing undertaken in an organisation, including both the technologies and their users. Ledley and Lusted are widely accredited for introducing computer systems into healthcare for the purposes of

supporting clinical decision making. Their seminal paper (Ledley and Lusted, 1959) outlined the use of statistical and probabilistic inferences to support differential diagnosis for clinicians. Their individual influences in the 1960s and 1970s were equally significant; as Ledley went on to invent the whole body CT scanner, while Lusted became a leading scholar in medical technology. Warner et al. (1961) published a mathematical model based on Lusted and Ledley's (Ledley and Lusted, 1959) work to diagnose congenital heart defects, which compared favourably when evaluated against experienced cardiologists. However, they noted the high sensitivity to false positive findings and errors in the system, and emphasised the need for accurate interpretation of clinical information. They advocated for an independent "gold standard" approach to judge the performance of the system. A few years later, Collen and colleagues (1964) developed a computerised differential diagnosis system at Kaiser Permanente that was based on Ledley and Lusted's analog computer card system. Later on, Bleich (1968) developed a system to suggest therapy for acid-base disorders, using clinical information such as arterial blood gases and vital signs. The system also had the ability to prompt the user where vital information was missing. It was unique in the sense that it added therapy to diagnosis by providing an evaluation summary and proposed management plan based on submitted information for review by the clinician. Wright and Sittig (2008) suggested that the algorithmic and closed nature of the clinical domain of acid-base disorders made it possible for such a system to succeed. These early systems were primarily viewed from a technical perspective, whereby the primary focus of their evaluation was on accuracy of diagnosis rather than the sociotechnical aspects.

Most of the systems that were developed in the 1970s followed the same model; whereby the user inputs factual clinical information, and the system suggests a therapy, diagnosis or

both. Evaluation of these systems focused on the technical capability of the system, appropriateness of the underlying rules compared with experienced clinical users and improvements in clinical performance. For example, in a controlled trial, de Dombal et al. (1972) investigated the merits of a computerised system based on a probabilistic model for diagnosing acute abdominal pain, and based on the analysis, establish the need for surgery. The authors concluded that despite the system being far from perfect, it matched the performance of senior clinicians and significantly improved decision making resulting in a 50% reduction in errors. The 1970s also saw the introduction of artificial intelligence, using expert system modelling techniques to develop systems to support diagnostic decision making. In 1974, Shortliffe (1975) developed MYCIN, an antibiotic prescribing system that used rules from its knowledge base to evaluate the clinical information provided and suggest optimal therapy. Initial evaluations of MYCIN suggested acceptable therapy in 75% of cases, which improved as more rules were added to the knowledge base. Miller (1983) developed the interactive ATTENDING system at Yale, which required the user to input clinical information and suggest a management plan. ATTENDING would then evaluate both the clinical data and the user's suggested plan, and then propose a plan for the user to review. This method was called "critiquing", and it led to "critiquing systems" that would later be used in the development of ventilator management, hypertension and other clinical domains (Miller, 1986). Miller et al. (1982) developed the INTERNIST-I system, which was a shift from previous systems that were characterised by limited functionality in narrow clinical areas. INTERNIST-I achieved reasonable success in its efforts to provide decision support across the vast spectrum of internal medicine using an expansive knowledge base. Despite its lack of commercial success, INTERNIST-I is credited for introducing the abstraction methods which are now widely used in complex diagnosis

based on the concepts of evoking strength, frequency and import functions. INTERNIST-I evolved into the Quick Medical Reference system, an expert consultant program that can also be used as an electronic textbook and intermediate level diagnostic spreadsheet derived from primary medical literature on diseases, diagnoses, findings, disease associations and lab information (Miller, 1989). Apache III was the first system to achieve commercial success. It was developed in the 1980s as a prognostic scoring system for decision support and quality assurance in intensive care units (Knaus et al., 1991). During the same period, Barnett et al. (1987) developed DXplain at Harvard Medical School. DXplain was an evolving diagnostic decision support system that explained its reasoning by using clinical findings to produce and rank possible diagnoses to explain the clinical presentation. DXplain used its extensive knowledge base to provide further advice on what clinical information to collect, while also highlighting usual and unusual characteristics of proposed diagnoses. A unique characteristic of DXplain was its user-friendly design that was suited for use even by physicians with no prior computer skills.

Most systems described in this section were developed in large medical and teaching centres in the USA. They were categorised as medical diagnostic decision support systems and were used by small teams of doctors within small patient groups. The majority of these systems were experimental in nature and were mostly limited to specific clinical areas. Their evaluation had a narrow focus, and was limited to validating their technical and clinical effectiveness. Most of the early legacy systems were confined to the clinical settings where they had been developed, and satisfactorily fulfilled local decision-making requirements but lacked the ability to transfer or exchange data within and across organisations. However, in the 1990s, there was a shift towards often-dispersed multi-disciplinary teams (Sittig, 1994). Also, the introduction of integrated care pathways



facilitated the development of systems to cater for the evolving needs of modern healthcare settings. These changes led to a shift in the classification and wider use of healthcare technologies and resulted in the introduction of clinical decision support systems to acknowledge the wider use of these systems beyond the small medical teams. However, some continue to refer to these systems interchangeably as health information systems, health information technology, medical informatics and lately e-health technologies (Bath, 2008). Liu and Wyatt (2011b) noted the ubiquity of health information systems in the delivery and management of healthcare services. They noted that although there were many arguments supporting their development and wider use, there was little evidence from the use of these systems based on structured evaluations. Most studies published in the 1990s were pilot studies, which although important, were not undertaken in real clinical environments and did not look at the long-term effects of these systems (Rigby et al., 2001; Kaplan, 2001b). Karsh et al. (2010) argued that these systems may not have the purported benefits on quality of care and costs. van Gemert-Pijnen et al. (2011) also noted the mismatch between postulated benefits and actual outcomes. Karsh et al. (2010) argued that success of these systems should not be measured on adoption or usage metrics, but instead on their impact on population health. Furthermore, they suggested that comparative effectiveness methods, such as the return on investment of each healthcare technology initiative compared with other alternatives would be a more appropriate evaluation strategy. The narrow focus of most evaluations and limited evaluation methodologies resulted in failure to identify the impact of CDSS on the broader healthcare systems (Karsh et al., 2010).

It has been noted that in practice, many health information systems are not evaluated at all (Talmon et al., 1999) and where evaluated, not adequately evaluated (Kaplan, 2001b;

Friedman and Wyatt, 2006). Wyatt and Keen (2001) also observed that although health information systems were often regarded as part of modernisation processes, they were left unevaluated for years following implementation. This, they argued, results in “technology lock in” and inflexibility. Furthermore, some systems which may initially be effective may be overtaken by the changing healthcare environment and thus require repeated evaluations (Wyatt and Spiegelhalter, 1990). Since the early 2000’s, there have been active discussions in health informatics and related disciplines focusing on the applicability and relevance of randomised controlled trials, which are strongly rooted in the medical tradition (Kaplan, 2001b; Ammenwerth et al., 2004; Heathfield et al., 1998; Wyatt and Spiegelhalter, 1990; Liu and Wyatt, 2011b). Proponents of the RCT argue that it is the most robust method to measure outcomes of health informatics applications (Friedman and Wyatt, 2006; Liu and Wyatt, 2011b). However, Kaplan (2001b) noted the lack of explicit theory to inform many CDSS evaluations, especially the “insulation” of CDSS evaluations from the wider informatics discipline. Also, these methods do not address the sociotechnical issues raised by Mumford (1983) or the organisational perspectives highlighted by Symons (1991) and Smithson and Hirschheim (Smithson and Hirschheim, 1998).

## 2.4 Evaluation of clinical decision support systems

The notion of clinical decision support systems (CDSS) is fairly new. In the past, CDSS have been referred to as medical decision support systems, with reference to the medical profession who were initially the main users of these systems. CDSS are part of the wider information systems. There is no widely accepted definition for CDSS. Some definitions narrowly and precisely relate to a specific clinical encounter, where a decision is made at the point of care (Hunt et al., 1998; Sim et al., 2001), while others broadly cover wider healthcare settings and multiple technologies (Osheroff et al., 2007). The diversity of CDSS definitions may also reflect the characteristics of the underlying technology being defined, and the clinical settings where it is intended for use. The following CDSS definitions have been selected to illustrate the different dimensions found in the literature:

Software that is designed to be a direct aid to clinical decision-making in which the characteristics of an individual patient are matched to a computerised clinical knowledge base, and patient-specific assessments or recommendations are then presented to the clinician and/or the patient for a decision.

Sim et al. (2001: 528)

Any electronic system designed to aid directly in clinical decision-making, in which characteristics of the individual patient are used to generate patient-specific assessments or recommendations that are then passed to the clinician for consideration.

Hunt et al. (1998: 1339-1340)

Provides clinicians, staff, patients or individuals with knowledge and person-specific information, intelligently filtered or presented at appropriate times to enhance health

and healthcare. It encompasses a variety of tools and interventions such as computerised alerts and reminders, clinical guidelines, order sets, patient data reports and dashboards, documentation templates, diagnostic support, and clinical workflow tools.

Osheroff et al. (2007: 141)

Sim et al.'s (2001) definition considers the individuality of the patient and specifies the function of the knowledge base, while also considering the patient as a potential decision-maker. In contrast, Hunt and colleagues' (1998) definition is more typical of a practitioner/patient relationship whereby the CDSS supports decision-making by the former. Osheroff et al.'s (2007) definition expands the role of CDSS to cover wider clinical processes and population health, and thus analogous with current trends in healthcare. However, as highlighted by Berlin et al. (2006), even seemingly specific definitions amount to naught if the complexities and heterogeneity of CDSS design, function and use, as well as host organization circumstances are not taken into account. As such CDSS definition depends on the type of CDSS, the intended use and users, and targeted population (Berlin et al., 2006). In the literature, CDSS is often used interchangeably with medical informatics, health informatics, health information systems, health information technology and other sub-categories to describe the various computerised health information technologies that are used to support healthcare professionals and systems (Bath, 2008). However, CDSS is more inclusive of other healthcare professionals besides physicians, who were the targeted user group of most systems up to the 1990s.

#### **2.4.1 Types and taxonomy of CDSS**

The early CDSSs such as MYCIN and INTERSIT-I were designed to simulate expert human thinking. These systems were adapted in the 1970s and 1980s to support real clinical

settings through knowledge based systems, i.e., assisting clinical decision-making rather than providing answers. The user would thus play an active part and interact with the systems, and select appropriate and discard inappropriate information, rather than passively receive outputs (Miller and Masarie, 1990). This was an important development that acknowledged the value of the decision-maker’s knowledge and expertise, consistent with findings from research on cognition and medical expertise (Klein and Calderwood, 1991; Miller, 1983; Miller, 1986; Miller, 1994; Miller and Masarie, 1990; Miller et al., 1982). According to Berner (2007), most CDSSs are comprised of the knowledge base, the inference or reasoning engine and a mechanism to communicate with the user using the steps in Figure 2.1.

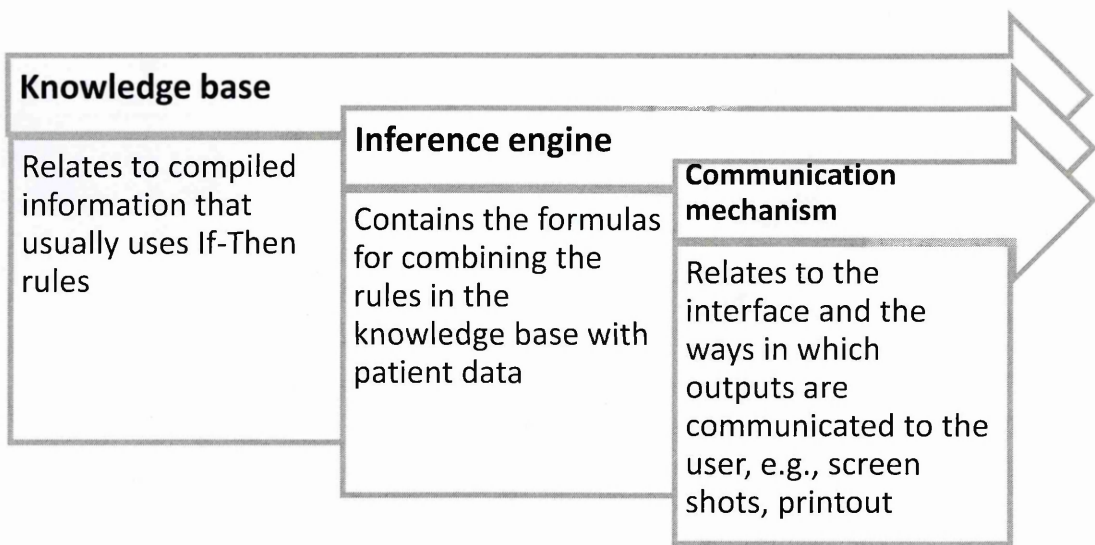


Figure 2.1 Structure of CDSS (adapted from Berner, (2007))

Using a taxonomy they had developed previously (Berlin et al., 2004), Berlin et al. (2006) described the characteristics of CDSSs reported in RCT-based studies between 1998 and 2003. The authors observed that CDSSs varied in design, function and use, context of use, knowledge and data sources, nature of decision support offered, information delivery, and workflow impact. They found that CDSS were widely heterogeneous and cautioned against

generalisation based on results from randomised controlled trials to different clinical or workflow settings. Their taxonomy was unique because it primarily focused on evaluation of CDSS. Other taxonomies have focused on technical issues (Kuperman et al., 1999) and information technology management perspectives (Teich, 1999) in line with trends in CDSS research. Berner (2007) also noted the diversity of systems that can potentially support clinical decision making, ranging from healthcare literature databases such as Medline, to tools incorporated in health information systems to perform financial and administrative analyses and sophisticated data mining techniques for both administrative and clinical data analyses. She argued that although retrospective approaches may be used in the development of guidelines, critical pathways or protocols used in clinical practice, they should not be considered CDSS.

The distinctions are important as vendors often will advertise that their product includes decision support capabilities, but that may refer to the retrospective type of systems, not those designed to assist clinicians at the point of care.

Berner (2007: 4)

A framework developed by Metzger, Perreault and colleagues (2002; 1999) distinguished CDSSs across three dimensions as follows:

- Timing of the decision support, i.e. before, during or after the clinical decision is made
- How active or passive the CDSS is, i.e., whether the CDSS actively provides alerts or passively responds to physician input or patient-specific information
- How easily CDSS are integrated into workflow processes

Berner (2007) noted that other distinctions between CDSS relate to whether they are standalone or integrated with existing patient records systems, and whether the decision support provided is general or specialty specific. She also noted the trend towards marketing of in-house developed (non-commercial) patient record systems, and incorporation of decision support functions into vendor developed computerised patient record systems and physician order entry systems to cater for the growing demand for CDSS. This section has shown that CDSSs vary significantly in their types and classifications. The heterogeneity has implications for CDSS evaluation because results of evaluations may not be easily generalisable to other clinical contexts. This is particularly important because randomised controlled trials and other experimental evaluation approaches which are favoured by many CDSS evaluators emphasise the importance of generalisability of such results into comparable areas.

#### **2.4.2 Benefits and uses of CDSS**

Bright et al. (2012: 2) described “classic CDSS” as “alerts, reminders, and order sets, drug dose calculations that automatically remind the clinician of a specific action, or care summary dashboards that provide performance feedback on quality indicators”. Other CDSSs are used for diagnosis, reminder systems for prevention, disease management, drug prescribing and dosing (Garg et al., 2005a). Classen and Bates (2011) reported that healthcare was “beginning to catch up” with other major industries in the adoption of information technologies. This followed a report by Hsiao et al. (2010) that almost 50% of USA outpatient practices and 44% of hospitals were now using electronic health records. Kaplan and Harris-Salamone (2009b) had also observed an increasing interest to invest in healthcare technologies in most developed countries aimed at benefiting from improved quality of healthcare and reduced costs. They noted other drivers including improved and

standardised processes of care, integration of healthcare systems and dealing with the burdens of an aging population. In the UK most initiatives for the adoption and implementation of healthcare technologies, and by inference CDSS are driven by the Department of Health and other arm's length bodies such as the National Institute for Health and Clinical Excellence (NICE) through clinical guidelines (NICE, 2010a). The literature shows that CDSSs have the potential to improve care (Kaplan, 2001b; Friedman and Wyatt, 2006; Black et al., 2011), particularly for systems designed for guideline adherence and those that aid physicians to make diagnostic decisions (Kaplan, 2001b). However, concerns have been raised regarding the fast rate at which new technologies are developed, often with little evidence of their effectiveness apart from early adopters' anecdotal accounts (Friedman and Wyatt, 2006). This also presents a challenge about how to evaluate existing and newly developed technologies. Ammenworth and colleagues (2004) argued that evaluating these systems should involve the interaction between the users and the relevant information processes, as well as the context in which they are used. However, these perspectives are often ignored in most evaluations and as noted by Kaplan (2001b), methods used in wider information systems' evaluation are little used in healthcare.

Some CDSSs are adopted for their potential to reduce errors through evidence-based medicine. Sim et al. (2001: 527) defined evidence-based medicine as "the practice of medicine based on the best available scientific evidence", and noted that its use with CDSS "therefore promises to substantially improve healthcare quality". The concept of evidence based medicine grew from efforts to standardise healthcare interventions in the 1980s. Ellwood (1988) called for "outcomes management", whereby all stakeholders would make "rational choices" using standards and guidelines and continually collecting, analysing and



utilising patient information to improve outcomes. This led to the development of guidelines to reduce variations in practice and improve the evaluation of healthcare services (Roper, 1988). However, (Schwartz and Mendelson, 1991) described the introduction of evidence based medicine as the “era of cost containment”. Harrison and Wood (2000) described evidence-based medicine as “scientific bureaucratic medicine”, and argued that this policy was designed to manage demand as a means to control healthcare costs in the NHS, rather than the purported clinical benefits. They argued that scientific bureaucratic medicine was based “on the assumptions that valid medical knowledge is derived from accumulated research evidence and that such knowledge should be implemented through clinical guidelines” (Harrison and Wood, 2000: 25). Clinical guidelines used in the NHS, such as the NICE guidelines are primarily based on evidence from randomised controlled trials, systematic reviews and other experimental methods. However, they have been criticised for often ignoring issues such as the potential effects of interventions such as changes in job roles and work redesign. For example, a study by Lomotan et al. (2012) highlighted that even within the specialist area of paediatric pulmonology, a guideline-based CDSS may not meet the needs of a small number of expert sub-specialists who have completely different skill-sets to their contemporaries in the same department. Other studies have found that physicians were non-compliant with guideline based CDSS (Durieux et al., 2000). Some studies have found that physicians poorly comply with guidelines, regardless of whether they are CDSS based or not (Trivedi et al., 2002; Bright et al., 2012). Some studies have also found little evidence whether CDSS assist physicians with diagnosis (Black et al., 2011; Bright et al., 2012). Patel and colleagues’ (Patel and Kaufman, 1998; Patel et al., 2002; Patel et al., 2000) studies of physicians’ cognition

concluded that although CDSSs helped to change their behaviour, it was unclear whether the thinking behind the behaviour had changed.

A systematic review by Garg et al. (2005a) found that CDSS improved practitioner performance in the majority of studies but patient outcomes remained “understudied and, when studied, inconsistent”. They also noted that improvements in practitioner performance were often associated with CDSS that prompted users in comparison with those that required users to activate the system. Additionally, they also found that evaluations undertaken by CDSS developers had higher incidence of practitioner improvement compared with those where the authors were not the CDSS developers. There have been efforts to standardise practice using CDSS following the Institute of Medicine (IOM) report (Kohn et al., 2000a) that highlighted the underuse, overuse and misuse of healthcare leading to variations and poor quality of care, and gaps in the translation of knowledge into practice in the USA. In light of these concerns, some CDSSs focused on improving care through changing clinicians’ behaviour through reminders and alerts at the point of care, implementing treatment plans and patient education (Kaplan, 2001b). Some CDSSs have been developed to reduce clinical errors (Bates et al., 2001; Bates and Gawande, 2003), and to improve efficiency and quality of care (Ammenwerth et al., 2004).

### **2.4.3 Approaches to CDSS evaluation**

Greenhalgh and Russell (2010a) observed that evaluations of health information technologies were largely based on “positivist”, “interpretivist” and “critical” philosophical assumptions. Positivist approaches take the view that an objective and measurable reality exists, and that such phenomena as project goals, outcomes, and formative feedback can be precisely and unambiguously defined (Orlikowski and Baroudi, 1991). Proponents of this

approach also argue that facts and values are clearly distinguishable, and it is possible to generalise the relationship between input and output variables (Friedman and Wyatt, 2006). They further argue that scientific testing is necessary to prove the worth of new drugs and healthcare interventions alike, including CDSS (Catwell and Sheikh, 2009; Friedman and Wyatt, 2006).

Health information systems should be evaluated with the same rigor as a new drug or treatment program, otherwise decisions about future deployments of ICT in the health sector may be determined by social, economic, and/or political circumstances, rather than by robust scientific evidence

Catwell and Sheikh (2009:1)

However, others (Heathfield et al., 1998; Kaplan, 2001b) argued that RCT-based CDSS evaluations are inappropriate because CDSSs are different from drugs and pharmaceutical products. Apart from the high costs of undertaking RCTs, they argued that the RCT method lacked external validity because results may not be relevant to a wider range of subjects or different contexts where CDSS usage may be different. Interpretivist approaches assume that reality is socially constructed, because people's perceptions on issues are different, and they assign different values and significance to facts (Klein and Myers, 1999). They argue that reality is never objectively or unproblematically knowable, and that the researcher's identity and values are inevitably implicated in the research process. Critical approaches assume that critical questioning can generate insights about power relations and interests within organisations (Klecun and Cornford, 2005). They argue that one purpose of critical evaluation is to ask questions on behalf of less powerful and potentially vulnerable groups, such as patients. The interpretivist and critical approaches view evaluation as "social practice", whereby one actively engages with the social situation, and considers the role of participants in framing and enacting with the situation (Pawson and Manzano-Santaella, 2012; Klecun and Cornford, 2005). According to Greenhalgh and Russell (2010a), interpretivist and critical approaches emphasise the need for "reflexivity", consciously thinking about issues such as values, perspectives, relationships, and trust.

They also refute the notion that rigour can only be achieved scientifically but instead, argue that the social factors can augment scientific considerations.

... rather, they hold that as well as the scientific agenda of factors, variables, and causal relationships, the evaluation must also embrace the emotions, values, and conflicts associated with a program...

Greenhalgh and Russell (2010a: 1)

Greenhalgh and Russell (2010a: 1) noted that while health information systems lie on the “technical” and “scientific” side, the “dreams, visions, policies, and programs have personal, social, political, and ideological components, and therefore typically prove fuzzy, slippery, and unstable when we seek to define and control them”.

Kushner (2002: 16) argued that although the positivist evaluation model is “elegant, in its simplicity, appealing for its rationality, reasonable in asking little more than that people do what they say they will do, and efficient in its economical definition of what data count”, it has several shortcomings. They argued that the multiple and often-contested goals of positivist evaluation, which leave it without a fixed goal that serves as a referent for comparison as its main shortcoming. An example is Greenhalgh and Russell’s (2010) study of the Summary Care Record in England, which found that the programme had numerous divergent goals.

Politicians were oriented to performance and efficiency targets, doctors saw the main goal as improving clinical quality in out-of-hours care, and civil liberties lobbyists perceived the program an attempt by the state to encroach on individual privacy

Greenhalgh and Russell (2010a)

Other shortcomings of the positivist model are instability of outcomes, unreliability of the causal link between process and outcome, indefinable characteristics of program success and counterproductive measurement of failure (Kushner, 2002). Greenhalgh and Russell (2010a: 2) argued that while experimental designs remain useful in many contexts, evaluation of health technologies required a paradigm shift to effectively deal with “politicised situations where goals and success criteria are contested”. Based on their experiences undertaking the Summary of Care Record study, the authors suggested a “set of guiding principles” for health technology evaluation based on social practice rather than scientific testing. The principles emphasise the need to consider the following aspects:

- Establishing the role of the researcher
- Setting out a governance process
- Allowing interpersonal and analytic space for dialogue
- Taking an emergent approach
- Understanding the dynamism of the macro-level context
- Understanding the dynamism of the meso-level context
- Considering the individuality of actors involved
- Consider the expectations and constraints inscribed in the technologies
- Using narrative as an analytic tool to synthesise findings
- Including critical events in relation to the evaluation itself

The authors argued that while positivist approaches remain largely relevant in clinical environments, they cannot ignore the “complex and fast moving socio-political arena”.

Differences in underlying philosophical position may lead to opposing quality criteria for “robust” evaluations. Some eHealth [and by extension CDSS] initiatives will lend themselves to scientific evaluation based mainly or even entirely on positivist

assumptions, but others, particularly those that are large-scale, complex, politically driven, and differently framed by different stakeholders, may require evaluators to reject these assumptions and apply alternative criteria for rigour. The precise balance between “scientific” and alternative approaches will depend on the nature and context of the program and probably cannot be stipulated in advance.

Greenhalgh and Russell (2010a: 4)

#### **2.4.4 Strategies for CDSS evaluation**

Kaplan (2001b) noted a tacit theoretical perspective in CDSS evaluation studies that leans towards quantitative methodologies and focus on measurable variables through comparison of CDSS with other alternatives. A systematic review by Ammenworth and de Keizer (2005) found that the majority of CDSS studies (79%) undertaken since 1982 were explanatory, and focused on testing a pre-defined hypothesis. Only 2% of the studies were exploratory studies focused on generating a hypothesis. The majority of exploratory studies evaluated organisational and social issues. During the same period, 83% (n=820) of studies used quantitative methods and only 5% (n=44) used qualitative methods. They noted that studies using combinations of quantitative and qualitative methods were increasing. The majority of exploratory studies used qualitative methods. Additionally, qualitative methods were mostly used for evaluating the organisational and social impacts of IT. In contrast with Kaplan’s (2001b) literature review, Ammenworth and de Keizer (2005) reported that since 1982, evaluation studies undertaken in clinical settings were steadily on the increase and fewer studies were now being undertaken in laboratory environments. They also found that most studies undertaken in the 1980s and 1990s primarily addressed a single evaluation focus, for example appropriateness of care, efficiency of work processes, software quality and quality of patient care. However, Rahimi and Vimarlund (2007: 398)

observed that CDSS evaluation studies “have been conducted with different aims, and their results are therefore difficult to compare”. Similarly, Klecun and Cornford (2005) highlighted some of the challenges faced by evaluators of health information systems, which they primarily attributed to the differences between the traditions of medicine and information systems.

The question of evaluation remains a major contentious issue, and perhaps even more so in the field of health informatics where the traditions of medicine meet and mingle with the information systems field... there is little agreement on the essential role of evaluation, a ‘best way’ to evaluate, on what and how to evaluate, whom to involve and within what paradigm to proceed.

Klecun and Cornford (2005: 229)

Klecun and Cornford (2005) suggested an evaluation approach broadly based on critical theory to highlight fundamental issues relating to the evaluation process and content. They argued that such an approach would also take into account “social, political and historical conditions” under which systems are developed and used. Additionally, power and political issues and structural constraints would also be addressed through reflective engagement and communication between stakeholders. They also acknowledged that the evaluation process could be constrained by vested interests, and the potential for evaluation to redistribute power (Klecun and Cornford, 2005).

Thus in the role of the researcher or evaluator (or both), we need to consider how our activity exposes, expresses, challenges or enforces some interests, while downplaying others, and to question our own alliances and assumptions.

Klecun and Cornford (2005)

2.4.5 CDSS evaluation methods

Concerns have been raised regarding that most health information systems were not being rigorously evaluated using methods applied in other safety-critical environments such as the airline and railway industries (Smith, 1992; Friedman and Wyatt, 1996; Friedman and Wyatt, 2006). Others have called for more experimental methods, particularly randomised controlled trials to measure the outcomes of CDSSs. Wyatt and Wyatt (2003) discussed the characteristics of three evaluation designs commonly used by healthcare organisations. Wyatt and Wyatt argued that due to the high costs of implementing large scale CDSS projects and their potential risk of harm to users and patients, the RCT was the best method to measure their effectiveness. Table 2.2 shows the typical evaluation methods that Wyatt and Wyatt noted are commonly used to evaluate health information technologies and the circumstances where they would be most appropriate.

	Simple before-after	Controlled before-after	Randomised trial
Typical use	Local audit	Regional decisions	National policy setting
Study role	To describe what happened	To suggest the cause	To determine the cause and size of the effect
Approximate minimum detectable effect	Large (>50% change)	Medium (>30% change)	Small (?10% change)
Chance of bias	Very high	Medium	Low if well designed
Scale	Within a single organisation	Within 2-5 organisations	The more organisations the better
Estimated lowest cost	Low	Low/Medium	Medium/High

Table 2.2 Characteristics of three major types of evaluation study designs [Adapted from Wyatt and Wyatt (2003)]



Ammenworth et al. (2004) noted that the timing of an evaluation was also an important indicator in deciding the appropriate evaluation method. They argued that appropriate evaluations should be undertaken throughout the lifecycle of the technology. They suggested that techniques such as software verification and validation would be essential during the development stage, while pilot and feasibility studies could be used for evaluating implementation. This could be followed by cost benefit and cost effectiveness studies. Lastly, monitoring studies could then be used to track how the system is working over longer periods of time.

#### **2.4.6 Pilot evaluation studies**

van Teijlingen and Hundley (2001) described pilot studies (also called feasibility studies) as smaller versions of a full scale study. These are important for pre-testing specific research instruments such as questionnaires or interview schedules to assess whether they work as intended and that further study is worth pursuing (Friedman and Wyatt, 2006; van Teijlingen and Hundley, 2001). Friedman and Wyatt (2006: 366-367) described feasibility studies as “proof of concept” evaluation that can inform decision makers whether a CDSS can be implemented and whether it can deliver the expected benefits. These studies also help to identify the immediate effects of an intervention and highlight where changes may be necessary. Pilot studies have also been used for patient safety oriented usability testing and validation of CDSS developed to offer automatic prompts to physicians to deliver treatment (Marcilly et al., 2012). However, van Teijlingen and Hundley (2001) noted that while a pilot study does not guarantee the success of the main study, it does increase the likelihood of success.

### 2.4.7 Before and after evaluation studies

Friedman and Wyatt (2006) defined before and after studies as comparative studies whereby a variable is measured during a baseline period and again after an intervention has been implemented. Measurements of the nature and frequency of the problem being studied can then be used to specify a new system or for later comparisons. Following implementation of a new system, measurements can then be taken to establish whether the problem has been resolved, as well as justifying the expenditure and judging the potential value of expenditure on future systems (Wyatt and Wyatt, 2003). They also noted that before and after studies can provide decision makers with valuable information. However, they argued that the 'after' results may be affected by other events going on in the organisation or human error, thereby making it difficult to attribute cause and effect. This potentially affects the quality of information available to decision makers.

“When studies generate the results we want, we take them at face value, and fail to register even major defects in study design. However, when they go against us, we launch investigations to find out why, and uncover all possible biases. Surely the rational approach is to start out with a more rigorous, bias free study design, so that whatever the result, we are satisfied with it and not left with lingering doubts about the real cause of the findings”

Wyatt and Wyatt (2003)

Friedman and Wyatt (2006) also highlighted the risk of bias and difficulties in reliably inferring cause and effect associated with before and after studies. They argued that this method causes the fallacy of the “after the event, therefore because of the event” (Friedman and Wyatt, 2006).

#### **2.4.8 Controlled before and after evaluation studies**

Controlled before and after studies are often used as an alternative to randomisation to improve attribution of cause and effect following introduction of an intervention (Friedman and Wyatt, 2006). This is done by introducing an external control, such as a similar organisation or patient group. However, Wyatt and Wyatt (2003) argued that even the task of finding a suitable external control may be challenging, especially collecting the pre and post implementation data. Additionally, the external control may not be as similar to the baseline as initially perceived, and may not be subject to the same factors. To counter this predicament, Wyatt and Wyatt (2003) recommended adding an internal control to strengthen the study by reducing bias.

#### **2.4.9 The Randomised Controlled Trial (RCT)**

The randomised controlled trial (RCT) is the primary method for testing new drugs and related pharmaceutical products (Kaplan, 2001b; Friedman and Wyatt, 2006). The RCT is an experimental study in which all factors that can potentially be manipulated by the investigator are controlled by randomly allocating participants to different study groups (Friedman and Wyatt, 2006). The RCT has also been used to evaluate interventions in other fields such as social care, education and criminal justice (Oakley et al., 2006; Zoritch et al., 1998). Liu and Wyatt (2011b) argued that the RCT was the only method that reliably allowed for the estimation of small but useful changes resulting from healthcare interventions, including CDSSs. However, Heathfield et al. (1998) noted that evaluating CDSS using RCTs had not provided significant evidence for improved patient outcomes or cost effectiveness. They also noted that RCTs are difficult to generalise and that their scope is limited, thus provide little detail for decision-making.

Liu and Wyatt (2011b) made a compelling case for the use of RCTs in evaluating CDSS outcomes on the basis that this method minimises bias and produces robust estimates of their impact. They argued that CDSSs that expose subjects to risk and those that come at a higher cost require robust evaluation using RCTs. Furthermore, they highlighted the shift towards evidence based policy making and need to demonstrate cost-effectiveness of all healthcare interventions. Garg et al. (2005a) also argued that claims that CDSS improve patient outcomes should be confirmed through clinical trials. However, Ammenworth et al. (2004) argued that the evaluation of CDSSs are often broader and more complex than would be the case when undertaking clinical trials or diagnostics studies. They noted that the primary focus of clinical trial designs is to objectively measure the impact of a therapeutic intervention on an individual patient. In contrast, evaluations of information systems measure quality, as well as effects on the structure, process and outcome of patient care. They also noted that the introduction of information systems takes time and the effects may not be immediately measurable. Others have also noted that users need time to get used to the systems and effectively exploit their benefits within their workflows (Palvia et al., 2001). Additionally, changes in work processes and staff, and modifications and additions to the system may result in changes in the intended uses, thus resulting in an often-changing evaluation target (Moehr et al., 2006). Additionally, where organisations implement a system in various departments, its usage and integration into workflow may differ in different settings. Ammenworth et al. (2004) argued that these changes and other factors make it challenging to control a typical healthcare environment in the same way as would be done in a clinical trial or to directly attribute causality between observed outcomes and interventions.

Lehmannm and Ohno-Machado (2011) noted that while RCTs were the generally accepted form of evaluation of healthcare outcomes, they were not always appropriate for ethical and economic reasons. They argued that emerging research methods such as comparative effectiveness were more relevant for evaluating CDSS. Comparative effectiveness research involves conducting new research or looking at available evidence of the benefits and harms of different interventions from existing research and provides information for decision makers (AHRQ, 2015). Lehmannm and Ohno-Machado (2011) noted that using comparative effectiveness research, evaluators could identify when RCTs were the most appropriate method or identify other suitable alternatives. Furthermore, they argued that most non-RCT studies did not expose patients, clinicians or other stakeholders to added risks like RCTs and were often less expensive to undertake. Examples include quantitative and qualitative evaluations and simulations of real clinical environments. Apart from documenting CDSS impact on clinical outcomes and processes, reports from these studies could also guide CDSS development as well as influencing future adoption decisions (Lehmannm and Ohno-Machado, 2011).

#### **2.4.10 Systematic Reviews**

Friedman and Wyatt (2006) described a systematic review as a secondary research method focused on answering a predefined question. It involves an exhaustive search and critical appraisal of relevant literature, extraction of data about methods and results and synthesis of results using appropriate methods. Proponents of RCTs argued that decision making by policy makers should primarily be based on evidence from RCTs and systematic reviews of relevant trials (Friedman and Wyatt, 2006; Liu and Wyatt, 2011b). However, Friedman and Wyatt (2006) noted systematic reviews were expensive to undertake and took much longer to conduct than commonly used evaluation methods. However, Wyatt and Wyatt (2003)

also argued that the cost of conducting systematic reviews were smaller compared with the potential costs of implementing an inappropriate system. Table 2.3 shows a range of CDSS systematic reviews that were carried out over the past two decades. Most of these systematic reviews were based on RCTs that were published in the English language. The majority of these reviews looked at the technical and clinical effectiveness of the CDSSs as well as improvements in the performance of the users. Very few RCTs looked at patient outcomes and a few reported improvements in patient outcomes. The CDSSs were found to be heterogeneous, and difficult to compare. Due to this heterogeneity, the majority of the studies used narrative synthesis, instead of pooling results or carrying out meta-analysis, thus making it difficult to generalise results. Consequently, most of the systematic reviews in Table 2.3 did not report any conclusive results and offered only limited recommendations for practice

Authors	Objectives of systematic review	Author's conclusions	Gaps in knowledge
<b>Montgomery and Fahey (1998)</b>	To assess the effect of CDSSs on the management of hypertension	CDSSs have a favourable effect on the uptake and follow up of patients in hypertension management	CDSS effects on user knowledge, information recording and blood pressure control was less conclusive
<b>Hunt et al. (1998)</b>	To assess CDSS effects on physician performance and patient outcomes	CDSSs can improve clinical performance especially for drug dosing, preventative care and other aspects of medical care but not convincingly for diagnosis	Effects of CDSS impact on patient outcomes had not been sufficiently studied. Healthcare organisations would also benefit from in-house evaluations before introducing CDSSs
<b>Shiffman et al. (1999)</b>	To assess the functionality and effectiveness of guideline-based CDSSs and their effect on changing user behaviours and improving patient outcomes	Authors' variations in types of guidelines, clinical settings and different systems, which made generalisation of results difficult. They also noted potential publication bias where	No meta-analysis due to variability of studies. The authors advocated for control of confounding variables to be able to adequately evaluate CDSS effects or to judge success or failure of CDSSs

favourable results were represented specific system implications			
<b>Shiffman et al. (2000)</b>	To assess the effects of guideline based handheld CDSSs in asthma clinics on adherence and patient outcomes	CDSS implementation resulted in increased costs of care without corresponding improvements in immediate and intermediate outcomes. Physicians resisted recommendations for guidelines which had not yet been validated	More research should be conducted before implementing guideline-based CDSSs in the management of asthma services for children. Context-specific costing and comparators would be required before implementing CDSSs
<b>Tan et al. (2005)</b>	To assess CDSS effects on the mortality and morbidity of new-born infants and the performance of physicians	Limited data from published RCTs made it difficult to assess CDSS effects	Authors recommended further RCT based studies assessing CDSS effects
<b>Kawamoto et al. (2005)</b>	To identify CDSS features deemed essential for improving clinical practice	The following CDSS features identified as essential:	The authors recommended that future research should be experimental, with focus upon the explicit description of interventions and the provision of details



	Automatic inclusion of CDSSs in clinical workflow	regarding clinician-system interaction.
	Provision of recommendations rather than just assessments	There may have been a bias towards those studies which showed significant findings for which data could be reliably extracted and overlooking findings from studies which were deemed to be ineffective
	Provision of decision support at the point of care	
	Provision of computer-based systems	
Garg et al. (2005a)	To assess the effects of CDSSs on practitioner performance and patient outcomes, and to identify the features of successful systems	By limiting to studies in English, the authors may have missed some key studies. The implications of findings on practice not specified. Authors recommended further research on patient outcomes.

<b>Kaushal et al. (2005)</b>	To review the effects CPOE CDSSs on medication safety	CPOE and isolated CDSSs significantly reduce medication errors and had benefits related to medication use	Further research recommended to look at commercial systems, the relative benefits of different CDSS types and factors related to their successful implementation
<b>Sintchenko et al. (2007)</b>	To determine the effects of CDSSs on patient outcomes	Improvements were noted in prescribing practices and treatment outcomes for patients with acute illnesses but were less effective in chronic conditions in primary care	Further research needed to quantify the range of CDSS benefits, explore the measurement of their effectiveness and usage in practice
<b>Randell et al. (2007)</b>	To assess the effectiveness of CDSSs on nursing performance and patient outcomes	CDSS introduction may not have positive outcomes	Further research required to identify those contexts in which CDSS use would be most effective

<b>Yourman et al. (2008)</b>	To describe the impact of CDSS interventions designed to improve the quality of medication prescribing in older adults	Various types of CDSSs may be effective in improving prescribing but data on clinical outcomes were limited	Future research should focus on developing better ways of measuring the effectiveness of CDSS interventions on improving clinical outcomes in older adults
<b>Wolfstadt et al. (2008)</b>	To assess the effects of CPOE with clinical decision support on the development of an adverse event	Few non-randomised studies had evaluated CPOE with clinical decision support in relation to the development of adverse events	Authors stated the need for further research to evaluate the effectiveness of CPOE with clinical decision support across various clinical settings.
<b>Schedlbauer et al. (2009)</b>	To assess the efficacy of electronic alerts and prompts on clinicians' prescribing behaviour	Most of the evidence demonstrated positive and often substantial effects of computerised prompts and alerts on prescribing behaviour	Authors recommended that additional RCTs of computerised alerts and prompts systems were required to determine the design features more strongly associated with positive effects on prescribing and clinical outcomes

<b>Kastener and Straus (2008)</b>	To evaluate the effectiveness of CDSS for management of osteoporosis	Multi-component tools for decision making targeting both physicians and patients may improve osteoporosis disease management	Further research needed for CDSSs for osteoporosis management. There was need for more rigorous evaluations to find the optimal disease management for osteoporosis
<b>Heselmans et al. (2009)</b>	To evaluate the effectiveness of CDSSs for implementation of clinical guidelines in ambulatory care (outpatient) settings	There was little evidence for the effectiveness of electronic guideline-based implementation systems	Future studies should evaluate patient outcomes (including long-term outcomes), potential harms and direct and indirect costs of guideline based CDSSs
<b>Eslami et al. (2009)</b>	To identify and characterise CDSSs and review their effects on the quality of tight glycaemic control process in critically ill patients	CDSSs improved quality control processes when measured by quality indicators. It was however impossible to define exact success factors because of lack of standardisation of glycaemic control, simultaneous implementation of	The authors recommended that future research should choose study designs that separate tight glycaemic control protocols from CDSS; investigate the most appropriate implementation site, target user and time of advice; and report

	CDSS with new treatment protocols and on usage, usability and satisfaction with various solutions used for integrating CDSS into users' workflow	
Pearson et al. (2009)	<p>To evaluate the impact of CDSSs on specific aspects of prescribing</p> <p>CDSSs appear to be effective for supporting some areas of the prescribing process, but there was little evidence for specific clinical domains and settings</p> <p>Authors recommended that more research on the short, medium and long term economic and clinical impacts of CDSSs, especially for stopping therapy. They also noted that different types of system and implementation strategies should be tested and the outcome measures should be clinical. Future studies should provide details of the individual features of the systems, their developmental methods, and the settings in which they were used</p>	

<b>Main et al. (2010)</b>	<p>To investigate which CDSS in order communication system (OCS) are in use within the UK and the impact of CDSS in OCS for diagnostic, screening or monitoring test ordering compared to OCS without CDSS. To determine what features of CDSS are associated with clinician or patient acceptance of CDSS in OCS and what is known about the cost-effectiveness of CDSS in diagnostic, screening or monitoring test OCS compared to OCS without CDSS.</p>	<p>CDSSs can have statistically significant benefit on process or practitioner performance outcomes in certain contexts.</p>	<p>Further research needs to be carried out to establish if there is evidence to determine whether CDSS in conjunction with communication order systems versus order systems alone for diagnostic screening or monitoring test ordering would be beneficial in the NHS context</p>
<b>Shojania et al. (2009)</b>	<p>To evaluate the effects on processes and outcomes of care attributable to on-screen computer reminders delivered to clinicians at the point of care</p>	<p>Point of care reminders generally achieve small to modest improvements in provider behaviour. A minority of interventions showed larger effects , but no specific reminder or contextual</p>	<p>Further research required to identify design features and contextual factors consistently associated with larger improvements in provider behaviour of</p>

	features were significantly associated with effect magnitude	computer reminders are to succeed on more than a trial and error basis
<b>Robertson et al. (2010)</b>	<p>To determine the impact of CDSSs targeting pharmacists on physician prescribing, clinical and patient outcomes</p> <p>Greater effectiveness found for safety focused compared with quality use of medicines focused CDSSs. The full benefits of quality use of medicines focused CDSSs may not be realised without good communication between pharmacists and physicians</p> <p>The authors stated that pharmacy based CDSSs needed to consider inter-professional relationships as well as computer system enhancements, Pharmacists outside institutional settings may require additional support to promote contact with physicians about appropriate medicines management strategies</p>	
<b>Bright et al. (2012)</b>	<p>To evaluate the effect of CDSS on clinical outcomes, healthcare processes, workload and efficiency, patient satisfaction, cost, and provider use and implementation</p> <p>Commercially and locally developed CDSSs were effective in improving healthcare process measures across diverse settings. Evidence for clinical, implementation</p> <p>The authors recommended that larger studies with longer evaluation duration were required to understand: how CDSSs might be expanded to accommodate multiple comorbid conditions</p>	

	economic, workload, and efficiency	simultaneously; how to determine which care team members should receive decision support what their effect was on clinical and economic outcomes; how they could be integrated into workflow across diverse settings; ways in which they might be incorporated into workflow tools
Nieuwlaet et al. (2011)	<p>To assess the effects of CDSSs for therapeutic drug monitoring and dosing on process of care and patient outcomes</p> <p>CDSSs had potential for improving process of care for therapeutic drug monitoring and dosing, but limitations in the evidence meant that no firm recommendation for specific systems could be made</p>	<p>Authors recommended that further research by independent researchers should be undertaken using cluster randomisation of non-specialised clinics and should focus on patient outcomes related to drug efficacy and safety</p>



<b>Sahota et al. (2011)</b>	To assess the effects of CDSSs on process of care and patients outcomes for acute medical care	Most CDSSs showed improvements in process of care, but patient outcomes were rarely assessed and were much less likely to show benefit. CDSSs for acute medical care have not matured sufficiently to enable decision makers to embrace technology for clinical application	The authors stated that high quality studies were needed to rigorously evaluate the effect of CDSSs on patient outcomes
<b>Roshanov et al. (2011)</b>	To assess the effects of CDSSs on chronic disease management processes and patient outcomes	Just over half of CDSSs showed improvements in the process of chronic disease management. Some improved outcomes. Evidence of effectiveness of computerised systems was limited	Authors stated that future trials with clear descriptions of the computerised systems, local context, implementation strategy, costs, adverse outcomes, user satisfaction and impact on workflow were needed
<b>Hemmens et al. (2011)</b>	To assess the effects of CDSSs for drug therapy management on process of care	CDSSs inconsistently improved process of care measures and seldom improved	Authors stated that future research should explicitly address patient

	<p>and patient outcomes, and to identify system and study characteristics that predicted benefit</p> <p>patient outcomes. Lack of clear patient benefit and lack of data on harms and costs preclude a recommendation to adopt CDSSs for drug therapy management</p> <p>outcomes, including potential harms, and costs and adverse clinician impacts of support systems. Future studies may wish to incorporate non-computerised approaches as active comparators</p>	
<b>Lobach et al. (2012)</b>	<p>To catalogue study designs used to assess the clinical effectiveness of CDSSs and knowledge management systems, to identify features that impact their success, to document their impact on outcomes, and to identify knowledge types that can be integrated into CDSSs/KMSs</p> <p>Strong evidence shows that CDSSs/KMSs are effective in improving care process measures across diverse settings using both commercially and locally developed systems. Evidence of CDSS effectiveness on clinical outcomes and costs and KMSs on any outcomes remains minimal. Nine features of CDSSs/KMSs that correlate with a successful impact of CDSS were newly identified or confirmed</p>	<p>Further research to assess the effectiveness of CDSSs on clinical outcomes and costs</p>

Hallin et al. (2011)	Do computer-based active (on screen) clinical decision support systems (CDSS) improve adherence to clinical guidelines?	Use of CDSS leads to some improvement in adherence to clinical guidelines for the studied outcomes: drug prescription, vaccination, and ordering of laboratory samples/tests. Patient-related outcomes have not been studied in this report. There is a vast variation in the magnitude of the reported improvement, ranging from no effect to a large improvement. Due to the vast variation, CDSS need to be evaluated separately. Regarding prescription or non-prescription of drugs, where deviations from the guidelines may cause serious consequences for the patients, there is moderate quality of evidence for a small to large improvement in adherence to	Further research need to assess how CDSS design, delivery, studied health problem, and outcomes influence the efficiency of CDSS as well as their cost-effectiveness
----------------------	---	--	--

clinical guidelines. Current knowledge about how CDSS design, delivery, studied health problem, and outcomes influence the efficiency of CDSS, is insufficient. Cost-effectiveness is unknown
---

Table 2.3 Summary of systematic reviews of CDSS evaluations (adapted from Cochrane database of systematic reviews)

#### **2.4.11 Formative evaluations**

The primary focus of formative evaluations is to provide adequate information aimed at improving decision making about the CDSS under study (Friedman and Wyatt, 2006). This is important because the development of CDSS requires a lot of planning and at times repeated rounds of prototype development and testing (Smith, 1993; Smith, 1998; Smith, 1992; Wyatt and Wyatt, 2003). Formative evaluations are essential for providing feedback or user comments to the development team, which can then be used to make the necessary revisions during the development stages (Ammenwerth et al., 2004).

#### **2.4.12 Summative evaluations**

According to Friedman and Wyatt (Friedman and Wyatt, 2006), summative evaluations help decision makers to determine whether the information technology has solved the problem it was implemented to solve. Summative evaluations are also important for documenting generic lessons regarding why a system has succeeded or failed, and also provide lessons for future adoption decision making (Wyatt and Wyatt, 2003).

#### **2.4.13 Monitoring as a form of evaluation**

Sittig (1994) argued that the increasing use of clinical software (including CDSSs) would entail that they required monitoring to assess their continued safety and clinical effectiveness. Friedman and Wyatt (2006) even called for the regulation of clinical software in the same way as medical devices. Regulators such as the UK Medicines and Healthcare Products Regulatory Agency (MHRA) and the USA Food and Drug Administration (FDA) have now developed guidelines for the evaluation of clinical software (MHRA, 2014).

#### 2.4.14 Frameworks for CDSS evaluation

The International Medical Informatics Association (IMIA) Working Group has held various workshops involving experts from various disciplines to address evaluation issues and promote evaluation as a methodology. The IMIA Working Group highlighted the lack of commitment to evaluation by decision makers in healthcare (Talmon et al., 1999). They sought to identify areas where they could make a difference, particularly by defining a methodology repository and developing a framework to help evaluators select the most appropriate methods. However, Rigby (2001) argued that although useful, developing methodologies and repositories alone would not achieve the required results. Instead, he argued that evaluations require adequate resourcing and commitment at national and local level. Furthermore, the “natural opposition to undertaking evaluation” should be directly addressed, especially by promoting the benefits of evaluations (Rigby, 2001: 1202).

Holle and Zahlman (1999) proposed a four phase model that includes a technical pilot study, feasibility study, controlled effectiveness study and cost effectiveness study. Talmon and et al. (1999) suggested the eight phase VATAM model, which looked at the entire technology lifecycle through the following stages: conception, design, development, integration, early use, exploitation, routine use and end life cycle. Other frameworks include the Health Informatics Systems Evaluation (HIS-EVAL) (Ammenwerth et al., 2004), the Guidelines for Good Evaluation Practice in Health Informatics (GEP-HI) (Nykanen et al., 2009) and the User-Centered Health Informatics Evaluation framework (Eisenstein et al., 2011). All three evaluation groups were made up of multidisciplinary teams who were tasked by the European Federation for Medical Informatics to develop and standardise evaluation guidelines for health informatics applications, covering topics ranging from study exploration to conclusion

(Nykanen et al., 2009). The User-Centered Health Informatics Evaluation framework recognised the need to adapt methods used in other fields and sought to revise economic evaluation, usability engineering and socio-technical analysis for use in health informatics evaluation (Eisenstein et al., 2011).

## **2.5 Contextual and organisational issues related to evaluations**

It has been noted that the context of CDSS use is often ignored by evaluations (Heathfield et al., 1998; Kaplan, 2001b). However, CDSS are complex social systems which affect and are affected by different actors, clinical settings and organisational environments where they are introduced (Greenhalgh and Russell, 2010a; Pawson and Tilley, 1997). Greenhalgh et al. (2010a) noted that it was essential to not only identify causal mechanisms that result in clinical outcomes, but also investigate and classify the contexts where CDSS are used. Kaplan (2001b) highlighted that most CDSS evaluations focus on isolated criteria or single outcomes and consequently fail to identify complex processes that contribute to the observed outcomes. The relationships between system characteristics, professional/individual/user characteristics, organisational characteristics and patient interests could also be explored as part of more comprehensive evaluations (Klecun and Cornford, 2005). Others have argued that CDSS evaluation should be more focused on organisational issues especially to support the organisational decision making processes (Lehmannm and Ohno-Machado, 2011). However, the context of CDSS evaluation differs between clinical areas and across organisations. This distinction is important because it affects how CDSS evaluation results are used or generalised across various settings (Greenhalgh and Russell, 2010a). Various

researchers have looked at organisational issues that affect the adoption and evaluation of healthcare technologies (Campion Jr et al., 2010; Kaplan, 2001b; Lluch, 2011). These range from existing organisational culture, structural issues and the intended users of the technology. Examples include the systems which are in place to enable technology adoption, social and professional issues and organisational readiness to adopt new technologies. Additionally, contextual issues that provide insight into how technologies are used and their effectiveness in specific clinical settings have brought new perspectives to both adoption and evaluation literature (Campion Jr et al., 2010; Lomotan et al., 2012).

A systematic review by van Germert-Pijnen et al. (2011) concluded that most healthcare technologies often fail to demonstrate sustainable success in clinical settings. They attributed these failures to the tendency to disregard the interdependencies between technology, human aspects and the socioeconomic environment, which often results in technologies that have a low impact on clinical settings. They also noted widespread failures in project management, minimal involvement of end-users, lack of project scope and definition, ownership and responsibilities of those involved. Chaudhry et al. (2006) also reported that healthcare professionals were often sceptical and hesitant to accept CDSS, arguing that they do not adequately support their work or result in patient benefits. van Germert-Pijnen et al. (2011) advocated for a holistic approach in the design and implementation of healthcare technologies that takes into account the complexity of healthcare and the rituals and habits of all stakeholders. They argued that such an approach would require careful coordination between various multi-disciplinary members involved, patients and other stakeholders. Kaplan (2001b) argued that the scope of CDSS evaluations should be widened to take into account social, organisational, professional and other contextual considerations. She argued



that unlike RCTs and other experimental designs, such evaluations would focus on the 'fit' of CDSS into the professional and organisational processes, rather than just focusing on measures of system performance and changes in physicians' behaviour. Similarly, Rahimi and Vimarlund (2007) noted that evaluation approaches such as user testing, cognitive studies, ethnography studies or socio-technical analyses would be more appropriate for assessing the effects of new systems on the organisation and its personnel. This also echoes the suggestion by Brooks et al. (2008) for shifting towards a more patient-centric approach to health informatics and their evaluations.

### **2.5.1 Existing technology infrastructure**

Existing technology infrastructure relates to the hardware, software, internal and external networks and related facilities that enable the adoption, use and evaluation of IT systems. Since the 1990s, information technology infrastructure has been identified as a critical factor in an organisation's ability to successfully adopt new technologies, optimise existing technological resources and competitively use IT resources (Snyder and Fields, 2006; Broadbent et al., 1999b; Broadbent et al., 1999a). Some models have been developed to assess the flexibility and responsiveness of an organisation's technology infrastructure. For example, Byrd and Turner (2000), explored the operationalisation of the concept of technology infrastructure flexibility and its value to the organisation. Chung et al. (2003) identified technology infrastructure flexibility as a core aspect of an organisation's ability to survive in fast changing business environments. They identified compatibility, connectivity, modularity and IT personnel as the key components that required strategic alignment to the implementation of IT resources. Others have explored the impact of technology enabled business process redesign (Broadbent et al., 1999b). Assessment of existing technology

infrastructure is essential is essential for CDSS evaluators because it may inform their evaluations as well as providing baselines to work from.

### **2.5.2 Organisational readiness for technology adoption**

Organisational readiness has been identified as an important factor for change management. Weiner (2009) described organisational readiness for change as a multi-level and multi-faceted construct that involves a shared commitment and collective belief of capability to implement change by the organisation's members. It involves an assessment of an organisation's readiness to adopt new ways of working, especially those that involve complex change. The assessment can be at individual, professional group, department or across the organisation, to establish the organisational capacity to work together towards the set objectives. Many strategies have been developed to assess organisational readiness for change involving collective behavioural change through systems redesign and involving multiple and simultaneous changes in work flow, staffing, decision making, communication and incentive systems (Holt et al., 2010). Weiner's (2009) theory of organisational change focuses on a thorough assessment of task demands, resource availability and situational factors of each change project. Weiner argued that when organisational readiness for change is high, its members are more likely to implement change, exert greater effort, exhibit greater persistence and display more cooperative behaviour, thus resulting in effective implementation and produce anticipated results. Using this model, implementation success depends on both change efficacy and contextual factors. Success is depended on the existing organisational culture and its members' collective belief that they have the capability to implement the change effectively. Savory and Fortune (2013) also highlighted the importance

of ensuring technology fit to the social context of the host organisation, existing organisational culture and the need for acceptance of the new technology.

### **2.5.3 External technology context**

The past two decades have seen a significant increase in the adoption of new technologies in the NHS and healthcare systems in most developed countries. Friedman and Wyatt (2006) observed that most technologies were small standalone systems that were not widely used beyond the clinical areas where they were developed. In the UK, the introduction of NICE guidelines and related national initiatives has led to increased demands for standardised care across the NHS. Some NICE guidelines have recommended computerisation of various aspects of care to deliver guideline adherent healthcare services (NICE, 2010a; NICE, 2007). This has put many NHS Trusts under immense pressure to find technological solutions to support their organisations' efforts in implementing guidelines.

## **2.6 Purposes of CDSS evaluations**

The purposes of CDSS evaluations vary considerably. Wyatt and Spiegelhalter (1991) grouped CDSS evaluation purposes into ethical, legal and intellectual categories. Others have grouped evaluation purposes into technical, clinical, economic, performance and regulatory categories among others (Bright et al., 2012). There have been calls for rigorous evaluations to support decision makers in making decisions about CDSS adoption (Ammenwerth et al., 2004; Karsh et al., 2010). Others have also highlighted the high costs of these systems and the need for decision makers to justify expenditure on them (Friedman and Wyatt, 2006) and comparing them with other alternatives (Karsh et al., 2010). Indeed, in one case, it took a computer crash for clinicians to find that it was quicker and easier to ask patients rather than use a

computerised system (Krakau and Fabian, 1999). Wyatt and Wyatt (2003) noted that decision makers in healthcare organisations are often pressurised by vendors, clinicians and the general public to develop and implement CDSSs. Others have found that decision-makers do not often have adequate information about these systems to make informed decisions about their adoption (Wyatt and Keen, 2001; Black et al., 2011; Lehmannm and Ohno-Machado, 2011). Wyatt and Wyatt (2003) noted that these systems are technology, rather than clinically driven and developers often do not see the need to evaluate these systems.

A systematic review by Kaplan (2001b) reported that most CDSS evaluation studies primarily focused on system accuracy rather than clinicians' performance when using these systems or assessing the impact of system use on patient care. Other studies have looked at physician perceptions of CDSS and others looked at the functionality of the CDSS and underlying knowledge base (Bright et al., 2012; Boonstra and Broekhuis, 2010). Heathfield et al. (1998) noted that CDSS evaluations were politicised and focused primarily on economic benefits rather than quality of life. They argued that consequently, most evaluations focused on justifying past expenditures and rebuilding trust with key stakeholders. Different stakeholders may have different priorities about which aspects of the system are evaluated. For example, Wyatt and Wyatt (2003) noted that organisational priorities for evaluating a new test ordering system would relate to whether the system saves money, reduce risk exposure or improve patient satisfaction. Clinical staff may be concerned about whether the system is easy to use, quicker than paper forms, clinical efficacy and effect on clinical decision making (clinical freedom). Concerns for laboratory staff would be whether tests ordered are appropriate and the volume of orders. Patients would be concerned about the time it takes for their tests and obtaining results. Heathfield et al. (1998) argued that CDSS evaluations should not only look

at accountability, but instead should aim to improve understanding of their role in healthcare. Furthermore, they argued that CDSSs should help to support the implementation of systems that deliver clinical and economic benefits.

This section has looked at the range of CDSS evaluation purposes found in the literature. The evidence suggests CDSS evaluations that look at a wide range of purposes are likely to provide useful information to decision makers than those with a narrow purpose (Heathfield et al., 1998; Kaplan, 2001b; Karsh et al., 2010).

## **2.7 Benefits of CDSS evaluations**

Despite the many challenges facing evaluators, Rigby noted that there were organisational, ethical and professional benefits of “rigorous evaluation studies” (Rigby, 2001). Many health information systems that were being implemented by organisations (including CDSSs) were novel in nature and involved significant changes in ways of working (Boonstra and Broekhuis, 2010; Rigby, 2001). Their introduction often involved many issues, such as hardware and associated devices and operational policies and procedures. Their novelty also exposes organisations and patients to potential risks (Wyatt and Wyatt, 2003; Ammenwerth et al., 2004; Karsh et al., 2010). Rigby (2001) argued that organisations had a moral imperative to generate and share knowledge about these systems through evaluations.

Others have argued that the scarcity of resources invested in healthcare systems should give further reasons for healthcare organisations to evaluate these systems and ensure their optimal and efficient use (Friedman and Wyatt, 2006; Rigby, 2001). Healthcare managers and policy makers have a duty of care to both the healthcare systems and clients they are responsible for and to account for their use of public money (Rigby, 2001; Friedman and

Wyatt, 2006). Rigby argued that this duty of care can only be ethically delivered through evidence-based approaches, which he argued, were scarce in health information systems. This paucity of evidence was attributed to the lack of investment in evaluation studies, which were largely based on false assumptions about the benefits of these systems (Karsh et al., 2010; Rigby, 2001; Rigby et al., 2001). Beauchamp and Childress (2001) discussed the responsibilities of healthcare organisations and healthcare professionals regarding their duty to do no harm to patients in their care. Another benefit of CDSS evaluation is to identify unanticipated, unwanted and harmful effects for patients [Karsh et al., 2010], and to reduce wastage of resources (Friedman and Wyatt, 2006). Rigby (2001) argued that it was imperative for organisations to monitor these consequences using evaluations and making the necessary changes as appropriate to ensure successful implementation.

There have been calls for healthcare policy and clinical practice to be evidence based, rather than anecdotal evidence or claims by suppliers about the efficacy of these systems (Friedman and Wyatt, 2006; Liu and Wyatt, 2011b). In the UK, regulatory organisations such as the MHRA and NICE have issued guidelines and standards based on the current evidence to classify and evaluate CDSSs (MHRA, 2014). NHS Trusts are required to audit their performance against these guidelines and standards. These audits are seen as important for standardising healthcare systems, promoting ethical practice and ensuring that avoidable adverse events do not occur (Rigby, 2001). Rigby noted that the development of business cases and technical specifications for CDSSs was often complex and time consuming. However, he noted that sometimes these processes were not done effectively, which in turn resulted in organisations repeating similar mistakes in the future. Rigby argued that organisations had an ethical responsibility to evaluate their systems to ensure that suboptimal processes were not

repeated in the future. Rigby also noted that the longer the health information systems were in use, the deeper there was organisational commitment to their continued use. He argued that in such situations, organisations would be wary of making any changes, however necessary that would be seen as disruptive to their continued use of existing technologies. There was therefore an ethical imperative to minimise operational disruption and resource wastage through structured evaluations that would help to identify the necessary changes and implement remedial actions timeously (Rigby, 2001).

There are many benefits of evaluating CDSS. They range from organisational, ethical, professional, economic and patient benefits amongst others. The literature suggests that CDSS evaluations that show clear benefits are likely to reduce barriers to evaluation and also likely to result in better decision making regarding their adoption and use.

## **2.8 Barriers to CDSS evaluations**

Barriers to CDSS evaluation include organisational, technological, human and methodological. Various CDSS evaluation studies suggest that there are disagreements regarding the role of evaluation, focus of evaluation and methodologies for evaluation studies (Kaplan, 2001b; Klecun and Cornford, 2005; Black et al., 2011). There are also questions regarding what can or cannot be evaluated (Burkle et al., 2001) and when, and how to evaluate (Friedman and Wyatt, 1996). Klecun and Cornford (2005: 230) also noted the controversy and challenges surrounding any evaluation activity, particularly establishing the context of evaluation, (“who is evaluating and why”), the process (“how the evaluation is done”) and the content (“what is evaluated”). Furthermore, they noted that CDSSs were often introduced simultaneously with other organisational change programs, thus making it difficult to isolate effects between

them. Kaplan and Shaw (2004: 217) noted that evaluation results may be viewed as site specific, and that the “methodological complexity of the undertaking, motivation, and ethical considerations” may lead others to abstain from conducting evaluation altogether or accepting evaluation results. Heathfield et al. (1998) noted the scarcity of credible evaluations of healthcare technologies. They argued that in part, this was a result of the assumptions that were made by decision makers that healthcare IT was beneficial on the basis of anecdotal evidence. Additionally, they noted that little or no credence was given to previous experiences or evidence from the evaluations in related disciplines (Heathfield et al., 1998).

The failure of CDSSs has often been attributed to the lack of involvement of clinicians in their development and implementation (Heathfield et al., 1998; Wyatt, 1994a; Wyatt, 1994b; Wyatt, 1994c). Heathfield et al. (1998) argued that many clinicians lacked the knowledge about evaluation issues to be able to assess the strengths and weaknesses of evaluation studies and meaningfully interpret their results. They noted that although a growing number of clinicians were now involved in the development and procurement of CDSSs, there was little useful information to support their decision making from evaluation studies. Most evaluations focused on economic and clinical outcomes, and ignored the fact that these outcomes were minimal and that results were difficult to transfer to other healthcare contexts (Heathfield et al., 1998). Ammenworth et al. (2004) identified the complexity of the evaluation object, the evaluation project and the motivation for evaluation as the three key problems facing evaluators of health information systems. They argued that evaluation approaches should endeavour to address these problems through the formulation of relevant questions, employing suitable methods and tools and utilising them appropriately. They



noted that evaluation problems could be attributed to the use of non-systematic study designs and poor management of the evaluation process. Klecun and Cornford (2005) also highlighted the increasingly complex nature of health information systems, both at technological and organisational levels. They noted that most systems involved various stakeholders and often spanned across professional and organisational boundaries, with wide-ranging potential to instigate changes in work practices and service delivery.

Ammenworth et al. (2004) argued that evaluation studies should look at various issues, such as the hardware or software, information processes and the interaction of health information technology with the users. They argued that evaluators needed to address the underlying social and behavioural processes, rather than just the technical and clinical aspects. Others have also noted the importance of issues such as CDSS integration with clinical workflow, implementation, the quality of information produced, training and ongoing support, the extent of usage and whether users were motivated to use the system (Berg, 1999; Palvia et al., 2001). Heathfield et al. (1998) highlighted the challenges involved in undertaking evaluations in real clinical environments. They noted that the main challenge was reconciling the perceptions and expectations of the various stakeholders involved, such as physicians, nurses, patients, administrators and funding agencies. They noted that stakeholders often had different perceptions about successful implementation of healthcare technologies, which could lead to multiple and possibly conflicting evaluation questions (Heathfield et al., 1998). Furthermore, many external influences such as regulatory authorities, legislation and macro-economic issues were also likely to compound the complexity of evaluation projects (Friedman and Wyatt, 2006; Ammenwerth et al., 2004). Tierney et al. (1990) noted that study designs from clinical trials could inform the evaluation of healthcare technologies.

Ammenworth and colleagues (Ammenwerth et al., 2004) also suggested the use of established traditions of designing, executing and reporting of healthcare technology studies based on clinical trials. However, others argued that simply adopting scientific approaches was not enough to resolve the challenges that were encountered when evaluating healthcare technologies (Heathfield et al., 1998; Kaplan, 2001b; Moehr et al., 2006). Furthermore, some studies found that many failed healthcare technology studies were not published, thus making it difficult to establish why healthcare technology evaluators faced so many challenges (Tierney et al., 1990; Ammenwerth et al., 2004; Karsh et al., 2010).

Ammenworth et al. (2004) noted the following barriers that affect CDSS evaluations:

- Unclear, conflicting or changing evaluation goals during study
- Large efforts needed for the preparation and execution of the study
- Complex and sometimes contradictory results
- Dependence of the evaluation results on the motivation and expectations of users
- Uncertainty whether evaluation results can be generalised to other environments

Wyatt and Wyatt (2003) presented the challenges in evaluating large health information systems. They noted that regardless of the potential benefits, some decision makers may argue that the resources could be better allocated elsewhere, for example on training and improving the quality of patient care. Kaplan and Shaw (2004: 216) described evaluation as inherently political, noting “what happens when a technology is introduced is affected by organisational and implementation processes, as well as affecting them”. Leys (2003) also noted the political nature of the “needs, values and interests” of various stakeholders affected by evaluation. May et al. (2003) also raised concerns that evaluations may be used for political justification to adopt or discontinue technology projects, rather than the

purported scientific grounds. Rigby (2001) highlighted many barriers to CDSS evaluation, which he argued were based on misconceptions about the effectiveness of CDSS. He noted that while the continuous development of CDSS was highly beneficial to practice, it was often influenced by subjective, “personalised demands and ideas” of individual professionals. He argued that this could be alleviated by using structured evaluation studies. However, he also cautioned that this could be seen as diversion of resources from “constructive uses”. He argued that evaluations could be disruptive and could also affect the professionalism, integrity and time awareness of those involved in the development and implementation of these systems (Rigby, 2001). Similarly, Wyatt and Wyatt (2003) looked at challenges faced by organisations when evaluating health information systems. They highlighted the complexity of the environments into which health information systems were introduced and sought to provide methodological solutions that would result in appropriate evaluations. Following a critique of commonly used evaluation methods, they concluded that experimental studies were most appropriate for evaluating large scale systems and where information was required for policy decisions.

A report by the NHS Confederation noted that healthcare organisations were struggling to effectively implement healthcare technologies because of top down initiatives and cultural barriers to their use (NHS Confederation, 2011). The report highlighted the challenges of introducing technologies into human systems, whereby there were multiple stakeholders with multiple interests. Boonstra and Broekhuis’s (2010) systematic review identified that adopting a change management perspective could potentially reduce barriers to the acceptance and usage of CDSS by physicians. They developed a taxonomy of barriers which were broadly categorised as financial, technical, time, psychological, social, legal,

organisational and change process. In particular, they noted that barriers were often interrelated and organisational and change process barriers were the key mediating factors for other barriers.

Liu and Wyatt (2011b) highlighted the increasing importance of evaluation in health policy formulation regarding expenditure on healthcare technology. However, despite the professed benefits of CDSS, studies looking at patient outcomes have not reported significant improvements [Kaplan, 2001]. Furthermore, many studies have found it difficult to establish which patient outcomes had been affected by CDSS introduction. According to Ammenworth et al. (2004), healthcare information systems were often associated with high costs, and amounted to nearly 5% of healthcare organisations' expenditure. The potential risk of their failure thus may have undesirable consequences on healthcare organisations, patients and users. However, healthcare organisations often fail to adequately assess the unwanted consequences of CDSS (Karsh et al., 2010; Kaplan and Harris-Salamone, 2009b). Karsh et al. (2010) highlighted many examples whereby CDSSs exposed patients and users to potential and actual harm. They argued that many decision makers made assumptions about the effectiveness of CDSSs based on anecdotal evidence rather than systematic evaluations. They noted that many CDSS failures remained largely unevaluated and underreported, despite their potentially harmful effects and cost implications. Additionally, they also noted that most studies generally reported positive results, and did not address unintended and unwanted effects, which may result in less informed decisions (Karsh et al., 2010).

This section has looked at barriers to the evaluation of CDSS in healthcare settings. Many of the barriers relate to the lack of systematic CDSS evaluations that take into account sociotechnical issues that are brought to bear on the introduction of CDSS into healthcare

organisations. These issues are difficult to address using randomised controlled trials and other experimental methods which are favoured by many CDSS evaluators. To reduce barriers to CDSS evaluation, it is necessary to consider the broader range of effects that come with their introduction into healthcare settings, particularly the human aspects and the context in which the evaluation is undertaken.

## **2.9 Trends in CDSS evaluation**

From the early legacy systems in the early 1970s, various methods have been used to evaluate CDSSs but none have emerged as providing a de facto standard. A review of CDSS evaluation trends by Kaplan and Shaw (2004) noted the inclusion of lessons learned and prescriptions for success in evaluation studies in the 1970s. The early 1980s saw an increasing interest in management issues, user satisfaction, and adoption and diffusion of information systems in medical informatics literature. They also cited the important links between various disciplines such as diffusion and communication studies, evaluation research and change management, noting how they shaped the development of medical informatics and health information systems in general. Various reviews of CDSS evaluation studies have been undertaken from different perspectives (Garg et al., 2005a; Kaplan, 2001b; Ammenwerth and De Keizer, 2005). However, despite their attempts to classify the subspecialty of interest, these studies tend to use terminology interchangeably and constantly refer to medical informatics, and by extension, medicine as the main field of study.

Kaplan (2001a) reviewed CDSS evaluation studies based on laboratory experiments and RCTs with a focus on changes to clinical practice and patient outcomes. The author also included 27 papers that are frequently referenced by other CDSS experts in her review. Kaplan

concluded that randomised controlled trials were the dominant approach used to evaluate CDSS, with a primary focus on changes in clinical or practitioner performance that could affect care. Using a different approach, Ammenwerth et al. (2004) investigated the trends of medical informatics evaluation studies from 1982 to 2002 with a focus on types of systems evaluated, evaluation methods, and country of origin. The authors found a significant rise in the number of evaluation studies in medical informatics between 1982 and 2002, but noted that explanatory research and quantitative methods remained dominant. They observed that the number of technically focused laboratory studies was declining and areas of increasing interest during this period were appropriateness and efficiency of patient care, user satisfaction and quality of software applications. They also noted an increase in studies that focused on quality of care processes and patient outcomes. They attributed this shift to maturation of medical informatics as a discipline. Ammenwerth and de Keizer's (2005) systematic review concluded that evaluation studies were increasingly considered important for the planning, development, implementation and operation of healthcare technologies. Friedman and Wyatt (2006) deemed the quality of evaluation studies as insufficient, while Garg et al. (2005a) noted that the methodological quality of CDSS evaluation studies had improved over time.

Delpierre et al. (2004) reviewed the impact of computer-based patient record systems from 2000 to 2003 on medical practice, quality of care, and user and patient satisfaction. They defined computer-based patient record systems as "computer software designed to be used by clinicians as a direct aid in clinical decision making", for example offering online advice, providing information or reminders to clinicians at the point of care. The authors reported a better understanding of the relationship between systems and medical practice, and noted

an increase in practitioner and patient satisfaction, which they argued, could lead to significant changes in medical practice. However, consistent with previous reviews, they noted that the “impact of computer-based patient record systems on medical practice and quality of care was not well demonstrated” (Delpierre et al., 2004: 414). They also identified other shortcomings such as lack of qualitative factors to explain the characteristics of the systems, disease and clinical setting where it developed and/or used, and stakeholder relationships, which may affect their use.

Computer-based patient record systems increased user and patient satisfaction, which might lead to significant improvements in medical care practice. However, the studies on the impact of computer-based patient record systems on patient outcomes and quality of care were not conclusive. Alternative approaches considering social, cultural, and organizational factors may be needed to evaluate the usefulness of computer-based patient record systems

Delpierre et al. (2004: 407)

Rahimi and Vimarlund (2007) reviewed the methodologies and findings of IT-based systems in medical informatics from 2003 to 2005. The authors observed that most studies focused on economic and organisational issues, and that results focused on “mostly positive outputs such as user satisfaction, financial benefits and improved organisational work” (Rahimi and Vimarlund, 2007: 397). They also noted the lack of evidence to support effective decision-making regarding evaluation and lack of a standardised framework to evaluate the impact of systems in clinical settings.

This review shows that there is no standard framework for evaluation effects and outputs of implementation and use of IT in the healthcare setting and that until today no studies explore the impact of IT on the healthcare systems' productivity and effectiveness.

Rahimi and Vimarlund (2007: 397)

A meta-level synthesis performed by Lau et al. (2010) reviewed health information systems evaluation studies published from 1994 to 2008 covering medication management, preventive care, health conditions, data quality and care process/outcome. They noted varying degrees of improvement of quality of care across the topics, notably high positive results in systems linked with guideline adherence such as immunisation and health screening, but no significant improvement in disease management and provider productivity. They concluded that key success factors influencing the success of CDSS included "having in-house systems, developers as users, integrated decision support and benchmark practices, and addressing such contextual issues as provider knowledge and perception, incentives, and legislation/policy" (Lau et al., 2010: 637). The majority of recent systematic reviews have concurred with findings from the studies discussed above, notably little benefit of CDSS to patient outcomes (Jaspers et al., 2011a; Bright et al., 2012), and significant improvement in practitioner performance (Jaspers et al., 2011a). However, Jaspers et al. (2011a) also noted that improvement in practitioner performance was markedly higher with drug ordering and preventative care reminder systems, because they required minimum patient data that were largely available before the decision support was generated. In contrast to Lau et al.'s (2010) review, Bright et al. (2012: 2) observed that both locally and commercially developed CDSS were effective at improving healthcare processes across diverse settings but concluded that evidence for "clinical, economic, workload, and efficiency outcomes remains sparse". Findings



and recommendations from various CDSS evaluation studies and systematic reviews have remained largely unchanged over the past decade.

The mismatch between health information systems research and practice highlighted by Coiera (2003; 2006) remains a central theme of most studies. Some have suggested that understanding how CDSSs are developed and what problems they are intended to resolve may be the key to this continued disconnect (Friedman and Wyatt, 2006; Kong et al., 2008). Kushniruk (2001) argued that findings from medical cognition and expertise studies provide important insights into the diversity of intended users of CDSSs in clinical settings, and the need to ensure that CDSSs are developed to cater for the needs of different stakeholders and their work settings. Studies by Klein and Calderwood (1991), Gaba and Howard (1995) and others in the 1990s set the trend towards studying reasoning, problem solving and decision-making in naturalistic settings, with practitioners making routine decisions in their work contexts. Leprohon and Patel (1995) investigated the relationship between decision-making strategies and the decision-maker's underlying knowledge. The authors found that when faced with high urgency situations, triage nurses in emergency departments resorted to simple rules that resulted in accurate decisions, but often-retrospective explanations of their actions did not correspond with decisions made. However, they noted that accuracy in the development of action plans and ability to assess the state of the presenting situation was evident in moderate to low urgency scenarios.

The development of various models and increasing research to better understand cognitive processes involved in decision-making in complex healthcare provides essential tools for the design and evaluation of CDSS and wider information systems in healthcare. As noted by Hammond (1998), decision-makers in high task complexity environments tend to simplify the

decision problem, using heuristic strategies based on previous experience and expertise. Patel and colleagues (2000) also argued for better understanding of cognitive process involved in decision-making, noting the existence of an inherent mismatch between the way humans and computers process information, with potentially detrimental effects upon clinical decision-making. Kushniruk (2001) noted that the majority of medical cognition studies have focused on diagnostic reasoning and problem solving, with particular focus on cognitive processes leading to treatment choice. He noted that most studies applied traditional approaches from judgment and decision literature, particularly the subjective expected utilities and decision outcomes theories. However, as highlighted by Miller (1994), most decision situations are often unstructured and semi-structured due to the individuality of patient presentation and complexity of disease classifications. Naturalistic studies by Klein and Calderwood (1991) and Gaba and Howard (1995) showed that expertise is an essential factor in complex decision-making under uncertainty, and that decision problems were mostly unstructured. They argued that using the expected utilities model as a foundation for technologies to support clinical decision-making would be tantamount to assuming that decision problems were structured and easily definable. However, as highlighted in previous sections, decisions that are made in clinical environments are fraught with uncertainties. Kong et al. (2008) noted that medical decision-making, reasoning and problem solving under uncertainty were becoming essential considerations for CDSS developers. They also highlighted the associated challenges of developing effective CDSS to support decision-making.

One of the main challenges in representation of and reasoning about medical knowledge is how to rationally handle those uncertainties so that a CDSS can support clinicians to make correct and reliable diagnosis and treatment decisions. Some identified issues include the representation of associated uncertainty in clinical domain knowledge, reasoning under uncertainty, the support of non-exclusive multi-part diagnosis and systematic clinical evaluation

Kong et al. (2008: 159)

This entails an understanding of what problem the CDSS is intended to resolve, how it resolves the problem, and under what circumstances might be valuable for CDSS evaluators (Friedman and Wyatt, 2006). In the next section, the key findings identified in this literature review will be discussed, followed by the research questions.

## **2.10 Key findings of the literature review**

As long ago as 1961 Warner et al. (1961) called for an independent “gold standard” approach to judge the performance of CDSSs. This led to the elevation of the RCT and other experimental designs to the top of the evidence hierarchy in healthcare evaluations (Cochrane, 1972; Friedman and Wyatt, 2006; Evans, 2003) but nevertheless the question of how best to evaluate CDSSs remains unresolved. This literature review has revealed that CDSS adoption in healthcare settings remains poor. Also where they were adopted, CDSSs were not being fully utilised to adequately support clinical decision-making. In some cases, CDSS were not evaluated at all. Kaplan’s (2001b) systematic literature review highlighted that decision makers in healthcare organisations were not getting enough information from evaluations to support CDSS adoption decisions. Kaplan (2001b) argued that this was because RCTs and other experimental methods that were used at the gold standard approach for evaluating

CDSSs were not geared to explain the low adoption and why CDSSs were not fully used or accepted in the clinical settings. The literature review has shown that the majority of systematic reviews of CDSS evaluations tend to include only RCTs and other experimental methods on the basis that they are more objective than qualitative and other none-experimental studies. However, most CDSS systematic reviews only include RCTs that were published in the English language only. This practice introduces publication bias, whereby evidence from unpublished papers and those in languages other than English are excluded. Also due to study and CDSS heterogeneity, most of the RCT-based studies use narrative synthesis instead of pooling findings together or conducting meta-analyses. Inevitably, this often leads to inconclusive results and sometimes cautious conclusions. Unsurprisingly, many CDSS studies do not explain how their findings impact on practice. As noted by Kaplan (2001b), this may suggest that the focus on RCTs and other experimental methods is too narrow and leads to researchers and decision-makers alike not gaining adequate information about CDSSs to help them to make informed decisions in practice. This literature review has identified five key factors of CDSS evaluation. In Figure 2.2, these factors have been used to construct a framework for CDSS evaluation that will now be examined.

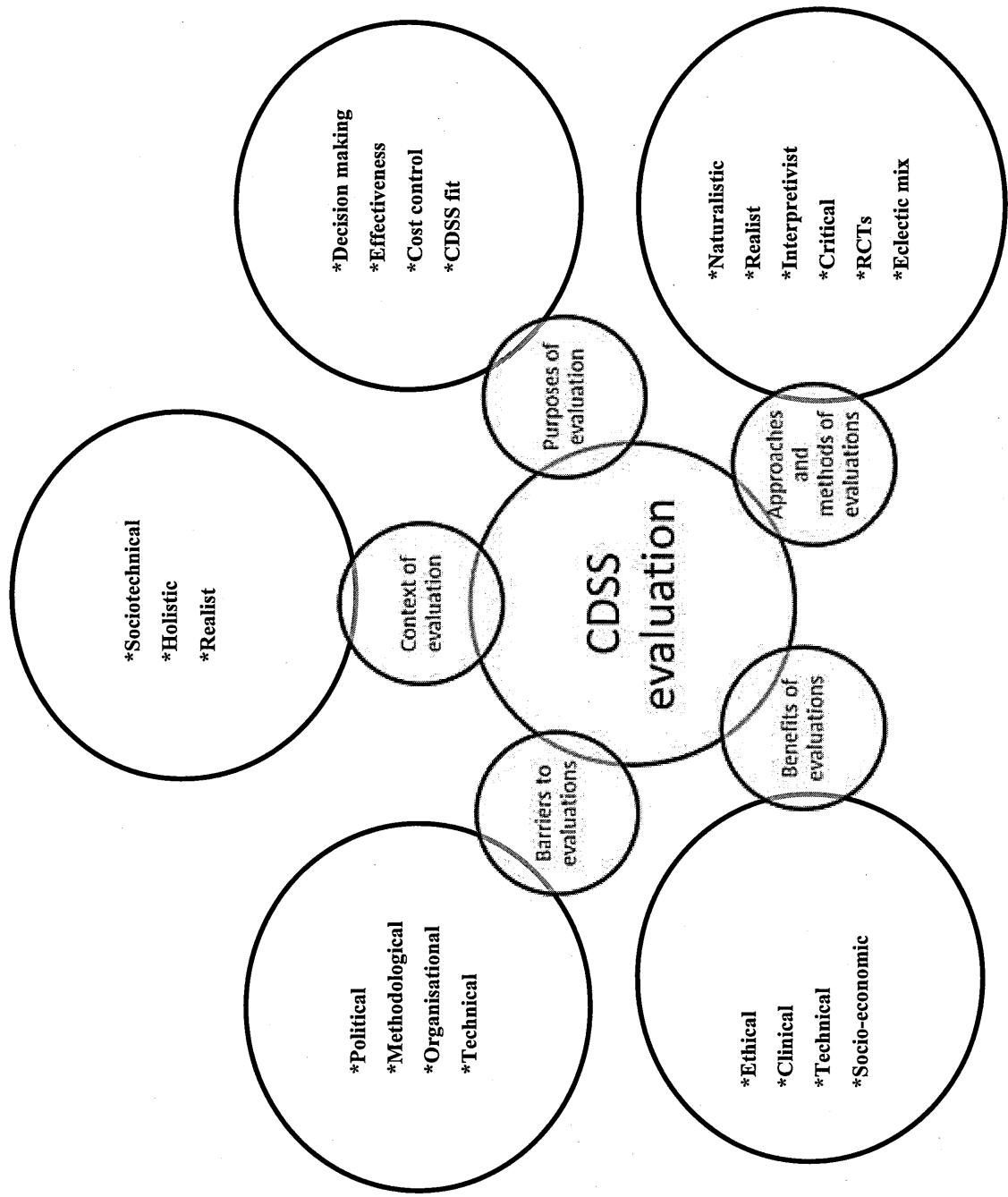


Figure 2.2 CDSS evaluation framework based on the key factors for CDSS evaluation identified in the literature review

### **Context of evaluation**

The context of CDSS evaluation has been identified as very important but the literature review revealed that it is largely ignored in published evaluation studies. This results in evaluators and decision makers failing to fully comprehend the complex interplay between CDSSs and the different actors and the organisational settings into which they are introduced. Lack of consideration for the evaluation context disregards the sociotechnical effects of CDSSs in environments where they are introduced.

### **Purposes of evaluation**

The literature review has shown that CDSS evaluation purposes vary considerably. The main purposes of CDSS evaluation identified in the literature review include technical and clinical efficacy; user performance and accuracy of clinical decision making; satisfying regulatory authorities; assessing cost effectiveness; and justification of IT investment to key stakeholders.

### **Approaches and methods**

CDSS evaluation approaches can be broadly classified as positivist, interpretivist and critical. Positivist methods focus on eliminating bias with a view to objectively and precisely establish and measure cause and effect. Interpretivist methods take into account the social construction of reality and contend that the researcher's identity and values are inevitably part of the research process. Critical approaches seek to question the status quo, with a view to gaining an understanding about power relations and the interests of the different stakeholders involved.

**Barriers to evaluation**

The literature review highlighted human, organisational, technological and methodological barriers to CDSS evaluation. Examples include disagreements regarding the role, focus and which methodologies that should be used for evaluation studies.

**Benefits of evaluation**

The literature review has identified many ethical, economic, professional and organisational benefits of evaluating CDSSs. Most CDSSs are novel and healthcare professionals and organisations have a moral duty to evaluate their effects in clinical environments and share this knowledge with the wider healthcare community.

Table 2.4 summarises that literature that supports this CDSS evaluation framework.

Key evaluation factor	Key Authors	Summary of CDSS evaluations factors
Context of evaluation	<p><i>Sociotechnical</i> - Mumford (1983);</p> <p><i>Holistic</i> - Kaplan (2001a); Kaplan (2001b);</p> <p><i>Realist</i> - Pawson and Tilley (1997); Greenhalgh et al. (2009)</p>	CDSS evaluation context is often ignored. However, CDSSs are complex social systems in themselves and their introduction into 'live' and often complex and established social systems is likely to affect (and be affected) by the various stakeholders. The context helps evaluators to not only find out what works, but "what works for whom in what circumstances and in what respect, and how?"
Purposes of evaluation	<p><i>Decision making</i> - Scriven (1972); Stufflebeam and Shinkfield (2007);</p> <p><i>Effectiveness of intervention</i> - Cochrane (1972); Evans (2003); Friedman and Wyatt (2005);</p> <p><i>Scientific Bureaucratic Medicine</i> - Harrison and Wood (2000);</p> <p><i>CDSS fit</i> - Kaplan (2001a); Ammenworth et al. (2004);</p>	<p>The novelty and cost of CDSSs requires adequate information from evaluations to inform decision makers and other key stakeholders.</p> <p>Evaluations should look beyond the achievement of set objectives of the CDSS project. Evaluation purposes should be widened to include the socio-economic, technical, clinical and sociotechnical effects of CDSS in the specific and wider clinical environments where they are implemented.</p>
Approaches and methods	<p><i>Naturalistic</i> - Guba and Lincoln (1989);</p> <p><i>Realist</i> - Pawson and Tilley (1997); Greenhalgh and Russell (2010);</p> <p><i>Interpretivist</i> - Kaplan (2001);</p> <p><i>Critical</i> - Klecun and Conford (2005);</p> <p><i>RCTs/Objectivist</i> - Liu and Wyatt (2011); Friedman and Wyatt</p> <p><i>Eclectic</i> - Stufflebeam and Shinkfield (2007)</p>	CDSS evaluations primarily use randomised controlled trials and other experimental methods. An eclectic mix of evaluation approaches and methods is required at every stage of the CDSS lifecycle. The choice of evaluation methods should be guided by the sets of research questions that need to be answered at the time and the needs of key stakeholders.



<b>Barriers to evaluation</b>	<i>Political</i> - Rigby (2001); <i>Methodological</i> - Guba and Lincoln (1989); Friedman and Wyatt (2005); Karsh et al (2010)	Barriers to CDSS evaluations include methodological, technical, human, political, social and organisational. Evaluation barriers should be minimised by employing a wide range of evaluation approaches that help to negotiate the political and ethical dilemmas, gaps, imperfections and often inconclusive deductions.
<b>Benefits of evaluation</b>	<i>Ethical</i> - Rigby (2001); <i>Clinical/Technical</i> - Healthfield et al. (1998); Friedman and Wyatt (2005), <i>Socio-economic</i> - Ward and Daniel (2006); Stufflebeam and Shinkfield (2007)	Benefits of CDSS evaluations should be aligned to the wider organisational goals throughout the CDSS lifecycle. Evaluation benefits should be a core part of the CDSS projects and should be communicated to all key stakeholders.

Table 2.4 Literature supporting the CDSS evaluation framework

This research project will aim to make a contribution to the CDSS evaluation literature across all five of the key factors that form the basis of the CDSS evaluation framework shown in Figure 2.2. In order to do this it will examine the following three research questions:

1. What are the key factors that affect CDSS evaluations in a typical NHS hospital setting?
2. How do these factors relate to the CDSS evaluation framework that was developed from the CDSS literature review?
3. To what extent do evaluations affect decisions to adopt CDSSs in healthcare settings and which evaluation methods are most likely to inform CDSS adoption decisions and why?

The next chapter will look at the methodology which will be used to carry out this research.

# Chapter 3 Methodology

This chapter sets out the process that was followed to investigate the evaluation of three clinical decision support systems (CDSSs) adopted by a UK National Health Service (NHS) Trust. In Section 3.1, a brief introduction to the research problem is given and the research questions outlined. In Section 3.2, the research strategies that were considered are discussed. The chosen research strategy is discussed in Section 3.3. Section 3.4 describes how the research was carried out, including the data collection and ethics approval processes. In Section 3.5, data analysis techniques within and across cases are outlined. The researcher's reflection on how the research was undertaken and the researcher's background and motivation to carry out this research is discussed in Section 3.6.

## 3.1 Introduction

The literature review reported in Chapter 2 has shown that despite major technological advancements and evidence of their potential to improve care, CDSSs are not widely adopted and where adopted, their acceptance and usage in clinical settings is low. Most CDSS evaluation studies use randomised controlled trials (RCTs) and other experimental methods that mainly focus on economic, technical and clinical effectiveness. Kaplan (2001b) noted that these evaluation methods do not address how and why CDSS are poorly adopted and reasons why they are not widely used in clinical settings. Others have also suggested that current evaluation methods may not be providing decision makers with adequate information to make successful adoption decisions (Bright et al., 2012). This research looked at three CDSSs adopted by a large teaching NHS Trust; a home grown system that was developed by a multi-disciplinary team of clinicians, ICT department developers and project managers and senior Trust managers; a bought-in system that

underwent significant customisation to the study Trust's clinical processes and another bought-in system with minimal customisation.

### **3.1.1 Research questions**

The research sought to answer the following questions that were developed from the key issues identified in the literature review:

- What are the key factors that affect CDSS evaluations in a typical NHS hospital setting?
- How do these factors relate to the CDSS evaluation framework that was developed from the CDSS literature review?
- To what extent do evaluations affect decisions to adopt CDSSs in healthcare settings and which evaluation methods are most likely to inform CDSS adoption decisions and why?

## **3.2 Research strategies considered**

Research strategies that were considered for this study include methods that are commonly used in healthcare research such as experimental, survey, archival analysis, history and case study as discussed by Yin (2009b) (See Table 3.1). Experimental methods such as the randomised controlled trial (RCT) focus on contemporary issues, particularly the “how” and “why” research questions and require control of the research variables. RCTs are widely used to measure the effectiveness of new drugs and other pharmaceutical interventions, cost effectiveness of treatments and interventions and measurement of clinical performance (Kaplan, 2001b; Greenhalgh and Russell, 2010b; Liu and Wyatt, 2011a). Proponents of RCTs argue that they are the most robust method to measure outcomes and eliminate bias (Liu and Wyatt, 2011a). Furthermore, they also argue that RCTs can be replicated in different settings and that results are generalisable (Friedman

and Wyatt, 2006; Liu and Wyatt, 2011a). However, others argue that this method’s reliance on controlling the research subjects and environment creates an artificial setting, which is often detached from real life clinical environments (Kaplan, 2001b; Greenhalgh et al., 2009). Additionally, the rigidity of the RCT design limits the researcher’s ability to take into consideration serendipitous circumstances in the course of research, such as unexpected data sources and changing research questions in light of new information (Creswell, 2007). The RCT design was not chosen because the researcher sought to observe the evaluation of CDSSs in their ‘natural environment’ of the hospital, with all its complexity and thus had no control of any variables.

	Form of research question	Requires control of behavioural events?	Focuses on contemporary events
Experiment	How, why?	Yes	Yes
Survey	Who, what, where, how many, how much?	No	Yes
Archival analysis	How, why?	No	Yes/No
History	How, why?	No	No
Case study	How, why?	No	Yes

Table 3.1 Relevant situations for different research methods [Adapted from Yin (2009b)]

Survey research is a quantitative method that involves the systematic collection of information from a selected sample to establish a representative picture of a large population (Ovretveit, 1998). According to Yin (2009b), survey research helps to answer research questions looking at “who”, “what”, “where”, “how many” and “how much” (see Table 1.0). Unlike RCTs, surveys do not require control of behavioural events. Other advantages of survey research include their low cost, efficiency and generalisability (Fowler Jr, 2013). However Fowler also noted that a poorly designed survey may result in poor sampling, which in turn may present inaccurate information to decision makers.

Additionally, poorly designed survey questions may result in inadequate answers and potential omissions. Furthermore, the sample may be distorted because of nonresponse caused by failure to contact respondents or their refusal to participate. Respondents may also fail to respond to some questions, thus potentially distorting the sample and generalisability of the results. Other limitations of surveys relate to their inability to generate new hypotheses by themselves or identify any variables that have been left out (Fowler Jr, 2013). The survey method was not chosen because of the potential limitations of the sampling techniques and its focus on quantitative variables. Furthermore, the survey method does not answer research questions relating to “how” and “why” research phenomena occurred as discussed by Yin (2009b).

Archival analysis is a method that involves the collection of data from archival records such as computer files, service and organisational records and previously completed surveys (Yin, 2009b). Archival records provide a static and permanent record that may otherwise not be found through other methods. However this method is time consuming, and records may not be available, may be incomplete or the researcher’s access to them may be restricted or denied. Archival analysis was not chosen as the primary research method because of potential challenges in accessing records due to commercial sensitivity and information governance issues. However, where the documents were made available, they were used to support the chosen research methods. According to Yin (2009b), history research method also deals with the “how” and “why” research questions and does not require control of behavioural events. Sources of evidence such as primary sources, eye witness accounts and documents are used to collect information and then write accounts of the past. The main disadvantage of history research is that primary data sources may be unavailable or incomplete and eye witness accounts inaccurate. Additionally, historical research deals only with events of the past. The historical research method was also not

chosen as the primary method because of its limitation to past events and the potential challenges in obtaining information from primary sources.

According to Yin (2009b), the case study method helps to answer research questions such as “how” and “why” (see Table 3.1). Unlike RCTs, case studies do not require control of behavioural events and are primarily conducted in the natural environment of the research problem. Case studies are widely used in psychology, medicine and nursing (Creswell, 2007). However, Flyvbjerg (2011) argued that case studies are generally held in low regard in academia or simply ignored and their “very status” as a scientific method is also questioned due to their perceived lack of reliability and validity. For example, Liu and Wyatt (2011a) argued that in evaluating informatics applications, case studies can only be used for exploratory studies and hypothesis generation due to their subjectivity. However, Eisenhardt (1989) discussed how case studies can also be used for systematic hypothesis testing and theory building. Flyvbjerg (2011: 302) noted that “much of what we know about the empirical world has been produced by case study research, and many of the most treasured classics in each discipline are case studies”. Walsham (1995) also noted the increase in the number of in-depth case studies that focused on human actions and interactions around the development and use of computer-based information systems in the 1990s. Leading journals such as MIS Quarterly have now adopted a more pluralist approach and now include qualitative studies in their publications.

### **3.3 Justification for chosen research strategy**

This research adopted an interpretive case study approach (Eisenhardt, 1989; Klein and Myers, 1999; Walsham, 1995). This was done to allow an in-depth look at the evaluations that were undertaken for each CDSS from the perspectives of the key actors involved. Walsham (2006) argued that our knowledge of reality and how we make sense of the

world is a social construction by human actors. It was therefore important to undertake the research within the settings where CDSS are used and collecting information from the key actors involved. Also, as noted by Yin (2009b) [see Table 3.1], interpretive case studies are the preferred method for answering “how” and “why” research questions, and in circumstances where the researcher has little control over events, or where the focus of the study is in a real life context. This research study met all the three circumstances cited by Yin and furthermore, the popularity of case studies in psychology and case analysis in medicine and nursing (see Creswell, (2007)) and the increasing importance of case studies in information systems research make it relevant to this study. Flyvbjerg (2011) noted the closeness of the case study method to real life situations and the wealth of details that can be acquired through the researcher’s access to multiple sources of evidence. He argued that this exposure helps to improve our understanding of human behaviour and also improve the researcher’s learning process and development of the necessary skills to undertake good research. Similarly, Yin (2009b) also noted that case studies are undertaken in natural environments, thus bringing the researcher closer to the research phenomena as they unfold in practice. Furthermore, case study method allows the involvement of more actors as the case develops, rather than restrictions imposed by quantitative methods such as randomisation. The importance of the researcher’s learning process through observation of events in their natural setting, regardless of the potential subjectivity and unreliability of case study research has also been highlighted (Campbell, 1975). Critics of the case study method argue that generalisations cannot be made from case studies (Liu and Wyatt, 2011a; Wyatt and Wyatt, 2003). However, Flyvbjerg (2011) noted that important scientific works such as Galileo’s rejection of Aristotle’s law of gravity, developments in physics by Newton and Einstein and much of Charles Darwin’s work were all based on case studies. He further argued that carefully chosen cases can



result in landmark scientific development and that formal generalisations are not the only sources of scientific knowledge. Others have suggested that the researcher's subjectivity could be limited by engaging with participants and clarifying issues during fieldwork (Creswell, 2007; Yin, 2009b). Eisenhardt (1989) also suggested choosing atypical or extreme cases to improve the richness of research data and internal validity. Flyvbjerg (2011: 311) argued that often it is easier to remember and make decisions based on meaningful stories from case studies rather than "meaningless data" from quantitative studies. He also noted that stories are easy to understand and act on than reading "meaning into data" from quantitative studies and then "making up stories". These reasons made interpretive case study the most appropriate method to undertake this research.

## **3.4 How the research was carried out**

This section describes how the research was conducted. The selection criteria and brief background to the three CDSSs is given in 3.4.1. Data collection methods are discussed in 3.4.2 and the processes of obtaining ethics approval described in 3.4.3. The first stage of the actual research involved the gathering of background information through informal discussions with key stakeholders in the study Trust to identify potential technologies and gain access. For all three CDSSs, relevant clinical pathways and policies, government policy and guidance from regulatory and professional bodies were reviewed. The background information that was collected and used to support the main research is outlined in Appendix 1.

### **3.4.1 Case selection**

The research was undertaken in one of the largest teaching NHS Trusts in the UK (see Chapter 3 for detailed Trust description). The study Trust was selected because of its size

and status as one of the largest teaching hospitals in the UK. It also has many similarities with various other comparable NHS Trusts serving similar socio-economic groups in other UK cities of similar size, thus will be representative to other NHS settings. The study was undertaken in a single Trust to maintain the same context for all three CDSSs. The cases were selected on the basis of being theoretically interesting as suggested by Eisenhardt (1989). Although all three cases involved CDSSs, they were atypical in their respective backgrounds and adoption circumstances. The selected CDSSs were similar in that they processed individual patient information and provided recommendations to the decision maker at the point of care. All three cases were based on NICE guidelines that were related to their individual clinical domains and additional guidance from relevant professional bodies. However, the CDSSs were on different parts of the internally versus externally developed CDSS continuum. T1 (see Chapter 4) was internally developed, while both T2 (Chapter 5) and T3 (Chapter 6) were externally developed. However, T3 underwent significant customisation to suit the study Trust's clinical work processes and was integrated with existing legacy systems from the outset. The study Trust's ICT department was involved in the development, implementation and evaluation of T1, and also worked closely with the developing company to implement, customise and monitor T3. However, they only had minimal involvement in the adoption, implementation and monitoring of T2. The research study was guided by Yin's (2009b) "embedded design", with "multiple units of analysis" in a single study. The embedded case design allowed each of the three CDSSs to be treated as a case on its own, within its own context, as well as the wider context of the study Trust. The study Trust was therefore the primary case, and the three CDSSs were the embedded units of the research study. The context of the study Trust was important because all three CDSSs were used within the same organisational environment, which helped to control some extraneous variations and limit the case

(Eisenhardt, 1989). The cross case analysis (see section 5.2) brought all three CDSSs together, in the context of the study Trust and other units of analysis such as professional groups and clinical departments.

### **3.4.2 Data collection**

Data were collected using a range of methods such as semi-structured interviews with key stakeholders, documentary analysis and observations (Yin, 2009b). Flexibility was maintained throughout the data collection process in order to take advantage of emergent issues and unique case features [see Eisenhardt (1989)]. The snowball sampling technique was used to identify the key individuals who were involved in the development, adoption, use and evaluation of the CDSSs and other sources of data as the case studies progressed (Biernacki and Waldorf, 1981). This technique allowed the researcher to follow up new leads to identify and interview new participants as the research developed.

#### **Direct observation (site visit)**

Observations included site visits whereby research participants or key stakeholders showed the researcher how the CDSSs were developed or adopted, how they were supposed to be used and a tour of some of the clinical environments where they were used. In all three cases, the researcher was guided by a senior departmental member who described the CDSS and demonstrated how it was intended to be used. The site visits provided an understanding of the individual CDSSs, the environment where they were being used and the background and supporting processes within the relevant clinical areas. Unlike interviews, the site visits provided unstructured observational data in the natural setting where the CDSS were used, which was essential to augment both the interview and documentary data as highlighted by Yin (2009b).

## **Semi-structured interviews**

Semi-structured interviews were carried out with key stakeholders, including senior managers and senior clinicians, clinical line managers, nurses and junior doctors. Semi-structured interviews were selected as the main data collection method to gain in-depth insights by drawing on information from a few key informants as recommended by Denzin and Lincoln (2009). These interviews allowed the researcher flexibility in terms of the order in which topics were considered, while also maintaining a clear list of issues to be addressed and questions to be answered. Additionally, these interviews had the added advantage of giving the interviewees opportunities to “develop their ideas and speak more widely on the issues raised by the researcher” (Denscombe, 2003: : 167). Interviewees were also given an opportunity to pursue their ideas or train of thought, and to use their own words, thus providing more in-depth information on themes or topics introduced by the researcher. Allowing the interviewee to “speak their mind” enabled the in-depth investigation of complex issues, particularly as personal accounts of experiences and feelings were important to this study. The interviews were conducted like “normal conversations”, but with a specific purpose and structure, starting with introducing questions regarding a key dimension of the respective CDSS, and allowing interviewees to give detailed descriptions of their experiences or understanding of the issue. This was followed by follow up questions on related issues to maintain the theme or establish causal links between issues as the interview progressed (Kvale, 1996a).

Denscombe (2003) noted that interviews rely mostly on the ability to conduct a conversation and do not require setting up complex equipment to collect data. These key informants were carefully selected based on their involvement in the adoption, use and evaluation of the selected CDSSs. Interviews were also appropriate for this study because

it involved the collection of information that could be considered as sensitive as noted by Denscombe (2003), for example power relations between various stakeholders in the study Trust so it was important to make participants feel at ease. As selected CDSSs were adopted and evaluated by a few individuals, there was value in conducting interviews with targeted stakeholders to access privileged information that the researcher would otherwise not be able to obtain anywhere else (Denscombe, 2003).

Most interviews were conducted on a one on one basis, in part because having one source of information made it fairly straightforward to locate specific ideas with individual participants. There was also flexibility to adjust research topics and questions during and between interviews as suggested by Eisenhardt (1989), as well as developing new lines of enquiry to clarify data accuracy and relevance at the point of collection. The interviews were conducted in a conversational, open ended style as suggested by Denscombe (2003). However, the research purpose and structure was maintained, starting with introducing questions regarding how participants first found out about the CDSS. This was generally followed by participants liberally giving detailed descriptions of their first experiences with the respective CDSSs, progressing to key contextual events which they experienced but often had not revisited. In most cases, several of the indicative questions were answered in the initial exchange, allowing the researcher to present follow up questions on related issues and establish causal links as necessary. The indicative interview questions were used to guide data collection and provide insights into interrelated and causal links. Examples of indicative interview questions that were used are provided in Appendix 2. These questions were used as a guide for the interviews but were adapted depending on the interviewee's role and responsibility within the organisation and the specific CDSS being considered. These questions evolved as the research progressed to explore new leads and were informed by previous interviews and documentary evidence.

(See Appendices 1.1 to 1.4.) The interview questions were structured to capture key intervals of the adoption stages as described by Friedman and Wyatt (2006) and key evaluation factors identified in the literature review (see Appendix 3).

Interviews were audio recorded with the consent of individual participants to provide “a more accurate rendition” and permanent record, which could be listened to again and again (Yin, 2009a: 109). Audio recording also allowed the interviewer an opportunity to focus on the interview dynamics, rather than writing notes, which could itself be a distraction during the interview and could result in incomplete responses (Kvale, 1996b). The researcher was mindful of the potentially disconcerting and threatening nature of audio recording and possibility of interviewees giving less honest answers (Walsham, 2006). To minimise this risk, all interviewees were given the opportunity to accept or decline, and to pause the recording in cases where they felt that issues under discussion were sensitive (Walsham, 2006). According to Kvale (1996a), audio recording also allows the interviewer an opportunity to concentrate on the topics and dynamics of interview, as well as offering a “decontextualized” version of the interview that can be re-listened to again and again. However, because interviews are a speech only record, field notes were also used to capture non-verbal and other relevant contextual factors to augment the audio record. All interviews were fully transcribed immediately after they were carried out. However, as the participants were specifically targeted key stakeholders, the need for accurate data was a key imperative. There was a possibility that this could not be achieved using shorthand notes, which may itself be a distraction during the course of the interview and may result in incomplete responses. Field notes were essential for noting the interviewer’s immediate impressions during the interviews, and capturing the contextual factors and non-verbal cues and relevant contextual factors as recommended by Eisenhardt (1989). Field notes are also less intrusive compared with audio recording.

This also helped to augment the audio record (Yin, 2009b). As noted by Yin (2009b), interviews may be affected by poor articulation of questions and answers, response bias, and deliberate or unintentional inaccuracies. Other common pitfalls of interviews include time required to carry out interviews, open format of data, interviewer effect and reliability issues, and inhibitions on the part of the interviewees (Yin, 2009b). To counter these pitfalls, interview questions were tested in pilot interviews in the initial stages of the research and refined to maintain a conversational style, while retaining specific focus as suggested by Denscombe (2003). Additionally, interview questions were adapted to operationalise evaluation concepts and suit participants' professional roles to allow the gathering of unique perspectives based on their experiences. As interviews are considered verbal reports (Yin, 2009b), they were corroborated with organisational documents related to the CDSSs, published and unpublished data and direct observations as discussed in succeeding sub-sections. By targeting key stakeholders, the researcher sought to gain an understanding of their priorities, opinions and ideas regarding how CDSSs were evaluated and how evaluations affected them and their work environments. Interviewees were given an opportunity to "expand their ideas, explain their views and identify what they regarded as crucial factors" (Denscombe, 2003: 189). The researcher also had the flexibility to adjust and develop new lines of enquiry as the interviews progressed, and opportunities to check the validity of data for accuracy and relevance at the point of collection. As interviews were prearranged with carefully selected respondents, the response rate was 100%. The researcher sought to build on the "personal element" offered by the interview method, whereby "people tend to enjoy the rather rare chance to talk about their ideas at length to a person whose purpose is to listen and note the ideas without being critical" (Denscombe, 2003: 190).

## **Documentary analysis**

Documents are an essential aspect of healthcare for clinical, professional, administrative and regulatory requirements. Yin (2009a) noted the usefulness of documents for strengthening the internal validity of case study research. This is achieved by augmenting other sources of evidence, triangulation purposes, and understanding communication channels between various professional groups. According to Yin (2009b), documents are a “stable, unobtrusive and exact” source of data that offers a broad coverage over time, which also provide insights into key events and settings. However, they can be difficult to retrieve or access, may be incomplete, and may contain reporting bias. Access to relevant documents was granted by the study Trust’s relevant senior managers and clinicians. Documents reviewed included letters, clinical audits, Trust reports and policies, clinical pathways and national guidelines, training materials, internal memorandum, email newsletters and Intranet communications, the study Trust’s annual reports and quality accounts for the relevant periods, and technology strategy briefings (see Appendix 1). Additionally, archival documents such as historic policy and service reports, lists of key people, press releases, previously undertaken surveys and audits were also accessed to gain fresh insights, triangulate evidence sources and improve internal validity (Yin, 2009b). Most archival documents were publicly available on the study Trust’s website and intranet and reports from regulatory bodies. Appendices 1.1 to 1.4 shows the range of relevant documents and related artefacts that were used to support the data collection process.

### **3.4.3 Ethics approval**

Before undertaking the research, ethics approval was sought from The Open University Human Research Ethics Committee (HREC). This process involved the submission of a research protocol, which included a literature review and research methodology, a



participant information sheet (PIS) detailing the purposes and conduct of the research, how data and participants' rights would be protected and an informed consent form (ICF) that was signed by both the researcher and each individual participant. The documents that were submitted for the HREC ethics application are provided as Appendices 4 and 5. Following HREC approval (see Appendix 6) all the documents listed above were submitted to the study Trust's Research and Development department. The researcher also had to complete the Gafrec guidance form as required for all research projects that involve collecting data within the NHS (see Appendix 7). The research project was approved by the study Trust's Research and Development department on condition that the research protocol was adhered to. Meetings were subsequently held with senior managers and clinicians from the relevant departments who confirmed their commitment to the project and approved access to the settings where the selected CDSSs had been adopted.

### **3.5 Data analysis techniques**

As recommended by Eisenhardt (1989), data collection and data analysis frequently overlapped to gain familiarity with the data and make preliminary assumptions at the point of collection. This process helped to speed up the data analysis process and allowed flexibility, such as changing data collection techniques where necessary. Changes included adding interview questions to probe new lines of enquiry and addressing new issues as they emerged (Eisenhardt, 1989). Corbin and Strauss's (1990) grounded theory approach was adopted to allow the researcher to continually immerse with the data to gain a deeper understanding of each case as the research developed.

### 3.5.1 Within case analysis

The case analysis adopted an approach based on Eisenhardt's (1989) within case analysis. This technique involves commencing data analysis during the data collection stage. It enables the researcher to manage large volumes of data collected from interviews and aides a detailed, descriptive write up of each case as recommended by Miles and Huberman (1984). Consequently, a "rich familiarity" (Eisenhardt, 1989: 540) can be established which helps to accelerate cross case analysis. Flyvberg (2011: 301) described this form of analysis as more detailed, complete and richer "depth for the unit of study than cross unit analysis". Within case analysis was important because it allowed the researcher to undertake "intensive analysis" of each CDSS. Field notes and running commentaries recorded during the research process were also an important part of the data analysis process (van Maanen, 1988). These notes covered observations, analysis and summaries of daily activities, such as emergent ideas, thoughts, initial impressions and perceptions were immediately recorded in a descriptive manner (Eisenhardt, 1989). Interview and documentary data were framed into the key themes identified from the literature review as suggested by Yin (2009b). This was followed by iterative analysis (Eisenhardt, 1989) of interview data using bullet points, spreadsheets and tabular displays, lists of issues according to interview respondents and professional groups and comparing extreme/polar themes (Miles and Huberman, 1984). The lists, spreadsheets and tables were also important because they allowed the researcher an opportunity to condense several pages of narrative interview data into manageable formats. Summaries of field notes were also compared with interview data and formal sources to confirm or refute the researcher's initial impressions. Documentary analysis of process maps and organisational structures were also performed to establish any causal links or gain fresh insights that may have been missed by the interviews (Yin, 2009b).

Using Yin's (2009b) "embedded multiple unit analysis", all three CDSSs landed themselves to three levels of analysis; the study Trust, the respective CDSSs, and professional levels. Data analysis mainly focused on the CDSS evaluations undertaken in each case, but also paid attention to organisational, professional and social factors as suggested by Kaplan (2001b) and Greenhalgh and Russell (2010b). Other considerations included external influences from the government and regulatory bodies such as NICE, current and previous evaluations of related systems within and outside the study Trust, previously undertaken clinical audits and surveys relating to the service supported by the CDSS. Roger's (1995) five stage adoption process (awareness, persuasion, decision, implementation, and confirmation stages) were used to provide structure to data collection and analysis. Each CDSS was written up using a descriptive framework [Lynd and Lynd, 1929] to identify causal links between different issues as outlined below [Yin, 2009]. The following format was maintained as discussed by Eisenhardt (1989):

- Descriptive background of each CDSS as described by the stakeholders and through documentary evidence
- Collating evidence from supporting documents and the environment where the CDSS was used, associated health information systems, non-computerised systems such as paper-based clinical pathways, protocols and other relevant sources of data
- Perceptions and expectations of the key stakeholders about CDSS evaluation, as well as their knowledge and experiences
- Assessment of the interactions and responsibilities between different stakeholders and professional groups in relation to CDSS usage and evaluation
- A narrative of how the CDSSs were evaluated at different stages of their adoption lifecycle

- Assessment of opportunities for learning from evaluations at practitioner/clinician, departmental and organisational levels
- Discussion and analysis of issues emerging from each study
- Cross study analysis of similarities and differences between identified themes

### **3.5.2 Cross case analysis**

Cross-case analysis commenced with detailed chronological descriptions of each CDSS. Summaries from each case were categorised using different formats such as lists, Microsoft Word® tables for direct comparisons to match patterns, establishing divergent issues and identifying frequency of issues (See Appendix 8). Comparable issues that were identified in the literature, such as perceptions of CDSS user groups, barriers to evaluation, evaluation methods, and level of decision support were used for cross-case categorisation. Lists of similarities and differences between the cases and unique insights from each data collection technique were used to corroborate evidence. These divergent techniques (Eisenhardt, 1989) were applied to analyse data beyond initial impressions during data collection and those from the within case analysis stage. Emergent issues such as CDSS effects on workflow patterns, contextual differences and influences from national initiatives such as clinical guidelines supporting individual CDSSs were also explored across the cases.

To conclude the cross case analysis, summaries of field notes and running commentaries for each case were compared for similarities and differences, taking advantage of the already established familiarity with each case as an individual entity (Eisenhardt, 1989). Key comparable dimensions across cases related to CDSS links with guidelines and the extent of such links, whether arbitrary or opportunistic; ownership of evaluation, for example whether centralised (NHS-wide or study Trust) or autonomous within the

respective departments or sub-specialist; resources made available for evaluation in relation to CDSS investment; stakeholders involved and the effect of clinical hierarchy or other organisational hierarchies; and the decision processes supported by the CDSSs. Timelines of external and internal activities and critical events were analysed case by case and across cases to enhance thematic analysis. Conflicting evidence was investigated to gain an understanding of the differences and establish previously unidentified patterns where applicable. Ensuing contradictory evidence was clarified with stakeholders and other relevant sources such as the literature and official reports.

### **3.6 Reflection on conduct of the research**

Undertaking this research project was an important learning curve. As a trained nurse and clinical researcher with prior experience working in various NHS clinical settings, I had a fairly good understanding of the NHS organisational and clinical structures. Throughout the 2000s, I experienced first-hand the transformation of various NHS services through advances in technology and related innovations ‘borrowed’ from other service industries and saw their mixed results and the effects they had on users and other stakeholders. I was always fascinated with technological innovations and wanted to further understand their wider implications beyond the clinical settings in which I was working.

Identifying a researchable question within the area of interest was challenging, particularly because of the limited methodologies used to study this topic and the novelty of CDSSs in the UK. However, qualitative studies of CDSSs by researchers such as Kaplan (Kaplan, 2001b), Greenhalgh and Russell (2010b) and others gave the researcher an opportunity to identify gaps in the literature. The researcher took the advice suggested by Eisenhardt (1989) to write field notes and running commentaries of the research as it progressed. Additionally, all the interviews were also fully transcribed by the researcher.

This process was very time consuming. However, it proved worthwhile because the need for accurate data from specifically targeted stakeholders was a key imperative for this study. Also as suggested by Eisenhardt (1989), field notes helped to develop meanings during data collection, and allowed comparisons with transcribed interview data during within case and cross case analysis. The early analysis also allowed flexibility with data collection such as responding to emerging issues by adding new interview questions [Harris and Sutton, 1986], excluding or including interview participants and taking advantage of new opportunities as they emerged to gain an in-depth understanding of each case (Eisenhardt, 1989; Gersick, 1989). Most interviews were performed on a one to one basis. However, on two occasions (T3), two participants were interviewed at the same time due to time pressures on their part. Although this was not planned, it resulted in the collection of in-depth data because the participants discussed their individual and collective perspectives of what they considered to be key events, and reminded each other of critical events in the adoption and evaluation processes. One of the immediate concerns was whether interviews would be audio recorded or summarized straight after the interview. The researcher was concerned that participants would feel uncomfortable with audio recording, especially since some of the issues discussed could potentially be viewed as sensitive. However, all interviewees consented to audio recording and this allowed the researcher to focus on the interviews and engage with the participants rather than writing notes during the interview. Interview participants were eased into the interviews by asking introductory questions regarding how they first came into contact with the respective CDSSs. All participants were willing to discuss their experiences and thoughts, mostly chronologically but in some cases the researcher had to refocus the interviews back to the indicative schedule. Many reflected on how the CDSSs affected them individually as well their teams and work streams.

## **Chapter 4 Background to study sites**

This chapter introduces the study site and describes the regulatory systems that affect the evaluation of CDSSs in NHS settings. In the UK, NHS organisations and healthcare professionals are regulated by numerous government and independent bodies that have different, overlapping and sometimes competing interests and expectations. It is necessary to discuss the effects of these bodies on the evaluations of CDSSs that are carried out by NHS organisations. Section 4.1 outlines the regulatory structure of healthcare in England, highlighting recent changes and the efforts to improve technology adoption and utilisation to support health service provision. Section 4.2 focuses on the regulation of health and social care professionals in England, particularly the role of professional and collegiate organisations in shaping and implementing government agenda. The study site (study Trust) is introduced in Section 4.3, noting its structure and technological landscape.

### **4.1 The regulatory structure of healthcare in England**

The Department of Health is a ministerial department that is responsible for developing policy and allocating funding for healthcare in England. To that end, twenty-three agencies and public bodies support it. It ensures the delivery and continuity of services and provides accountability to parliament, the public and taxpayers. The key areas covered by the Department of Health include public health, the National Health Service (NHS), social care and public safety emergencies (see Figure 4.1). Its priorities for 2014/15 include addressing the challenges of an ageing population through reform of social care and integrating services for older people; improving standards in healthcare delivery; utilising technologies to improve the quality of healthcare and planning for a more sustainable healthcare system.

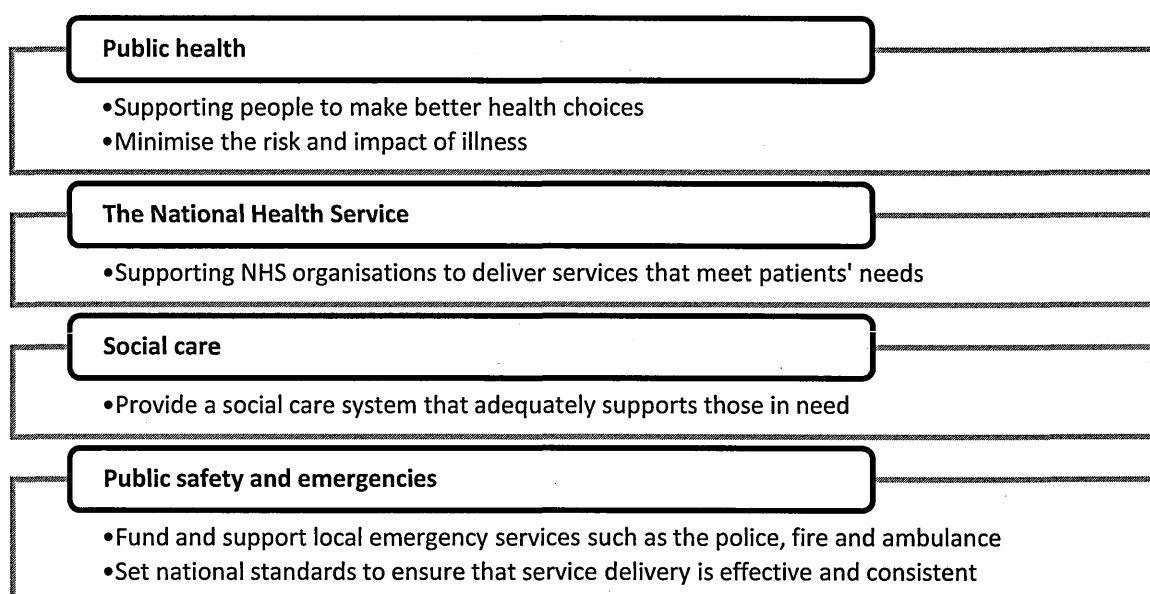


Figure 4.1 Key functions of the Department of Health

The Department of Health White Paper, *The new NHS: modern, dependable* (Department of Health, 1997) highlighted variations in the quality of care provided by NHS Trusts. It also noted the slow response to the Bristol Royal Infirmary case (Teasdale, 2002), which had resulted in “serious lapses in quality”. These failures led to the establishment of the Commission for Healthcare Improvement to offer guidance to NHS organisations on Clinical Governance. In 2003, the Healthcare Commission was established to combine the Commission for Healthcare Improvement and the National Standards Commission, which was an independent body responsible for the inspection and regulation of residential and domiciliary care. The Healthcare Commission also worked closely with the Audit Commission to improve efficiency in NHS organisations through assessment of their performance against national standards, inspecting NHS providers and recommending special measures where performance was found to be poor. The Healthcare Commission was accountable to parliament and commissioners, and published annual reports on the



state of NHS services. The use of audits to assess conformity to national standards has remained a key method of evaluating NHS services.

NHS England is a department of the Department of Health responsible for funding and delivery of NHS services and improving health outcomes for people in England. It works closely with the Clinical Commissioning Groups (CCGs) and associated organisations and NHS providers to improve decision-making and quality of healthcare. It also gives advice to patients about their health and lifestyle through the NHS Choices website. Additionally, NHS England also sets the agenda for technology, systems and data management in the NHS through the following streams:

- Strategic systems and technology
- Data and information
- Information governance
- The care.data programme (better information means better care)
- Care.data advisory group

Through the Strategic Systems and Technology Directorate, NHS England formulates policy for NHS information technology and informatics to guide commissioners, providers and suppliers to make informed investment decisions, identify alternative approaches and deliver the highest quality of care to patients. This Directorate also aims to support the development of systems that fit existing clinical workflows and enables healthcare professionals to effectively undertake their roles as well as benefitting patients and the public. The key deliverables of the Strategic Systems and Technology Directorate include facilitating the adoption of innovative and safe methods of record keeping, enabling access to patient records across specialty and institutional boundaries, enabling and supporting individuals to access their health records online, launching the Choose and

Book service and making eReferrals available to patients and healthcare professionals by 2015. The Technology Directorate also sought to support hospitals to implement systems for safe and effective prescribing for patients and making integrated digital care records available to users at the point of care. Another important role of the Technology Directorate is the commissioning of national information technology infrastructure to support NHS services such as the national Spine System, which enables the sharing of information across NHS services, the N3 Network and NHSmail, which is a secure email service. Although NHS organisations are responsible for developing their individual technology strategies, changes in priorities at national level may affect local initiatives in both positive and negative ways. National priorities may also dictate technology evaluations that are undertaken at local level.

#### **4.1.1 Parliamentary regulation of healthcare**

The Health Select Committee is one of nineteen Select Committees which were established by the House of Commons under the Standing Order No. 152 to work with Government Departments. The Health Select Committee oversees the operations of the Department of Health and its associated bodies through examination of policy, administration and expenditure and any other relevant subjects on enquiry. The Health Select Committee has the power to send for “persons, papers and records”, and can insist upon the attendance of witnesses, production of papers and other relevant materials. The UK parliament currently has more than 500 All-Party Parliamentary Groups that are concerned about diverse issues such as animal welfare and civil liberties and venous thromboembolism. These groups are comprised of cross party Members of Parliament and peers from the House of Lords. To ensure formal recognition by the House of Commons, each group must have twenty members at all times from both the governing

party and opposition. The all-party groups often act as pressure groups for their selected issues and help to keep the government, the opposition and Members of Parliament informed of parliamentary and outside opinion. The All-Party Parliamentary Thrombosis Group was set up in 2006 to promote the awareness of venous thromboembolism (blood clots) amongst parliamentarians. Their role expanded to venous thromboembolism prevention nationally, focusing on spreading awareness, assessment, management and prevention of venous thromboembolism. The group publishes research reports and hosts headline events and advocates for best practice in venous thromboembolism prevention across the NHS. It worked closely with the Department of Health to implement the National Venous Thromboembolism Prevention Programme (see Chapter 5). They also survey all acute NHS Trusts in England against national standards for venous thromboembolism prevention and identify areas requiring improvement. For example, in 2013, each NHS Trust was given an individualised scorecard, outlining its compliance and allowing them to benchmark their performance against other NHS Trusts (See Chapter 5).

#### **4.1.2 Independent regulatory bodies**

A number of independent regulatory bodies work closely with the Department of Health to regulate, develop and implement policies and guidelines across the NHS. The National Institute for Clinical Excellence (NICE) was originally set up in 1999. Its aim was to assess and approve the most clinically and cost effective drugs and treatments for use in NHS Trusts in England and Wales. In 2005, NICE merged with the Health Development Agency, forming the National Institute of Health and Care Excellence. Its mandate expanded to developing public health guidance to help prevent ill health and promote healthier lifestyles. In 2013, NICE became a Non Departmental Public Body, and its responsibility extended to developing guidelines and standards in health and social care for England as

set out in the Health and Social Care Act 2012. Although sponsored by the Department of Health, NICE is operationally independent of the government. Independent committees of experts make NICE guidelines and recommendations and its Board and Senior Management Team are responsible for developing its strategic policies and operational decision-making. Guideline development committees primarily use evidence from randomised controlled trials and systematic reviews because these methods are deemed to provide the highest quality of evidence. By August 2014, NICE had developed 850 guidelines covering a diverse range of topics such as venous thromboembolism, acutely ill patients in hospitals, prostate cancer and medicines adherence among others.

NICE has seven Directorates covering clinical practice, public health, health technology evaluation, communications, health and social care, evidence resources and business, planning and resources. The Centre for Health Technology Evaluation develops guidance and technology appraisals on the use of new and existing treatments and procedures in the NHS, such as medicines, medical devices, diagnostic techniques, surgical procedures and other interventions. Its research and development team is responsible for developing and improving methods used in guideline development and commissioning relevant research. The Health and Social Care Directorate is responsible for improving quality in the NHS through quality standards and their implementation in practice. It is also responsible for NICE Pathways, which are online tools that combine all NICE guidelines, quality standards and other related materials into easily accessible formats. The Health Technologies Adoption programme also falls under the Health and Social Care Directorate. It facilitates the adoption of selected medical and diagnostic technologies in the NHS. The Evidence Resources Directorate manages databases that provide authoritative evidence and best practice relating to new medicines in development. This Directorate is also responsible for the Information Management and Technology and the Information

Resources teams. These teams are responsible for NICE digital services and identifying, selecting and appraising new evidence. Although NICE Pathways are supposed to be used within the context of individual NHS Trusts and their existing initiatives, in most cases, they are applied as de facto “how to” guides in clinical areas. This may have implications for the evaluations that are carried out by NHS Trusts because they may primarily focus on issues that are receiving the most attention at national level, while missing useful information on relevant contextual issues.

The Health and Social Care Information Centre is part of the Department of Health Informatics Directorate, which replaced the NHS Connecting for Health in March 2013 [HSCIC, 2014]. Its primary role is to be the authoritative source of data and information relating to health and social care. It also supports the delivery of IT infrastructure, information systems and standards to improve patient outcomes. Its catalogue of data includes official statistics, results from surveys, audits and reports that are collected from various sources in health and social systems. The Health and Social Care Information Centre also produces guidance from hospital based information including clinical audits and data quality resources for clinicians, Hospital Episode Statistics, Patient Reported Outcome Measures, Secondary Uses Service, and the Summary Hospital Mortality Indicator (The Health and Social Care Information Centre, 2015).

Another important regulatory body is the Medicines and Healthcare Products Regulatory Agency (MHRA). This is an executive agency of the Department of Health responsible for regulating medicines in the UK. The MHRA also produces guidance on medical devices and stand-alone medical software, as well as outlining requirements for CE marking for stand-alone software that is used as a medical device. Stand-alone software (software medical device) is defined as software which has a medical purpose at the time of it being placed

onto the market (MHRA, 2014). It does not include software that is incorporated into an existing medical device, such as software that controls the function of a heart scanner, which is deemed to be part of the device. However, the regulation of software medical devices is limited by the intended purpose as defined by the manufacturer. The Medical Device Directive defines a software medical device as “software... intended by the manufacturer to be used for human beings for the purpose of: diagnosis, prevention, monitoring, treatment or alleviation of disease, diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap, investigation, replacement or modification of the anatomy or of a physiological process, control of conception..... (MHRA, 2014). T2 and T3 (see Chapter 6 and 7 respectively) are both registered with the MHRA. Before and after registration of a software medical device, developers are required to carry out various evaluations throughout the lifecycle of the software medical device’s development and implementation, focusing primarily on the technical and clinical efficacy, as well as post implementation longitudinal data collection to show that the device is safe for use in clinical areas. However, it is up to individual developers to decide whether or not their device requires registration, and thus effectively may have an effect on which evaluations are performed, if at all.

Dr Foster Intelligence was launched in 2006 as a joint venture with the Department of Health. Its aim is to improve the quality of health and social care by monitoring the performance of the NHS and providing healthcare information to the public. Dr Foster Intelligence works collaboratively with various stakeholders such as NHS organisations, the Department of Health, local government, academia and the private sector. Its Dr Foster Unit at Imperial College London is responsible for developing methodologies that facilitate the identification of potential problems in clinical performance as well as highlighting areas of achievement. Dr Foster Intelligence also helps providers and

commissioners to benchmark the quality and efficiency of health services, and cost and clinical effectiveness against key indicators. Examples include the Dr Foster Care Quality Tracker which collates latest data from multiple sources and is linked with the Care Quality Commission's Hospital Intelligent Monitoring indicators. The Care Quality Tracker is an "early warning system" that enables NHS Trusts to timeously identify and investigate alerts before notifications are raised by the regulators. NHS Trusts can also monitor and measure their quality outcomes and patient safety through Dr Foster's Real Time Monitoring tool and the hospital standardised mortality ratios (HSMR). Dr Foster Intelligence affects all three CDSSs selected for this research through HSMR monitoring and its proximity to the Department of Health, and specifically the VTE risk assessment tool because Dr Foster undertook audits looking at how individual NHS Trusts were implementing government policy (see Chapter 5). Dr Foster's influential position has an impact on evaluations that are carried out by NHS Trusts because they are required to submit information on their performance monthly and in some cases, they are ranked according to the results. However, some NHS Trusts (including the study site) have questioned the credibility of DR Foster Intelligence methods of collecting information and presenting results. They argued that these methods failed to take into consideration other key performance criteria and relevant contextual issues.

Another key Department of Health body is the National Patient Safety Agency (NPSA). The NPSA aims to identify and reduce risks, and improve the safety of care provided to patients by NHS organisations in England and Wales. Confidential reports on patient safety incidents are reported by NHS organisations through the National Reporting and Learning System. These reports are then analysed by clinicians and patient safety experts to identify risks and opportunities to improve patient safety and provide feedback and guidance as necessary. The NPSA works collaboratively with the Royal Medical Colleges,

NHS staff and related organisations, patient groups, the Department of Health and its agencies, academia and other stakeholders. The NPSA developed a root cause analysis methodology, which seeks to identify systemic and process failures in clinical areas, learn from them and implement action plans to ensure that they do not recur. The NPSA root cause analysis methodology is widely used across the study Trust, particularly by the Clinical Governance and Audit and Effectiveness departments in collaboration with departmental Clinical Directors and Nurse Managers as a de facto evaluation method for investigating serious adverse events. This method was also adopted by the study Trust's Thrombosis Committee to investigate all cases of venous thromboembolisms that occurred within 90 days of a hospital admission (See Chapter 5). Results from these investigations were shared across the study Trust's clinical specialties to learn from failures and improve processes where recommended.

## **4.2 Professional regulation and collegiate societies**

Healthcare professionals in the UK are regulated by a number of bodies such as the General Medical Council (GMC) and the Nursing and Midwifery Council (NMC) among others. The GMC and NMC are independent regulators responsible for registering doctors and nurses respectively in the UK. Their key roles include; keeping up to date registers of doctors and nurses respectively; fostering good practice and promoting high standards of relevant education and training. The Academy of Medical Royal Colleges brings together expertise from its members and Specialty Faculties to improve health and patient care through education, training and setting quality standards. Its work is primarily commissioned by the Department of Health and crosses boundaries between NHS organisations and the Royal Colleges and Faculties and primarily focuses on the following areas:



- Promoting high standards of education, training and professional development of its members
- Revalidation, governance and administration
- Undertaking quality improvement initiatives in collaboration with the Department of Health and other related bodies
- Advising and helping the government to develop and implement health policies
- Monitoring standards and promoting evidence based practice through clinical guidelines and performance audits
- Representing members and engaging with key stakeholders

The regulatory organisations (GMC and NMC) and Royal Medical Colleges supported the government initiatives that led to the development and implementation of all three CDSSs that are included in this research. In addition there was also support from a number of specialty collegiate societies such as the British Society for Haematology (see Chapter 5), the British Society of Urological Surgeons (see Chapter 6), and the Intensive Care Society (see Chapter 7).

The British Society for Haematology is the main professional body for Haematologists in the United Kingdom with about 1500 members. Its primary role is to advance the practice and study of haematology and bring together people who are interested in the subject. Its official journal is the British Journal of Haematology. It also publishes guidelines from the British Committee for Standards in Haematology, which provides Haematologists with up to date information on the diagnosis and treatment of haematological conditions. The Society actively supported the Department of Health's VTE Prevention programme which resulted in the development and implementation of the T1 (see Chapter 5). The British Association of Urological Surgeons is a professional organisation for urological surgeons

which aims is to promote high standards of practice in urology to improve patient outcomes through education, research and promotion of clinical excellence and audits. It supported the development and implementation of T2, as well as encouraging NHS Trusts to adopt the system and facilitating the necessary funding and commissioning (See Chapter 6). The Intensive Care Society is the main professional body for intensive care professionals in the UK. It aims to deliver high quality critical care services to patients through the production of guidelines and standards, organising national meetings, training courses and focus groups. It also represents the interests of intensive care in the Royal Colleges and the Department of Health and other related organisations. The Intensive Care Society's APACHE II study in the early 1990s compared intensive care units in the UK and found significant variations in outcomes (Rowan et al., 1993). This led to the setting up of the Intensive Care National Audit and Research Centre (ICNARC) in 1994 to provide an independent national resource for the monitoring and evaluation of intensive care through its National Audit Programme. Also in 1994, the ICNARC Case Mix Programme was set up to audit patient outcomes from adult, general critical care units in England, Wales and Northern Ireland. The ICNARC Clinical Trials Unit designs and conducts research to improve the care and outcomes of critically ill patients using randomised controlled trials, observational studies, systematic reviews and methodological studies. The Intensive Care Society was instrumental in the adoption of early warning scoring systems that led to the development of T3 (see Chapter 7). The National Cardiac Arrest Audit (NCAA) was jointly set up by ICNARC and the Resuscitation Council (UK) in 2008 for in-hospital cardiac arrest from all acute UK hospitals. Data collection started in October 2009 and comparisons of outcomes following resuscitation are now published in the NCAA reports. The aims of the ICNARC audit are to:

- improve patient outcomes;

- decrease incidence of avoidable cardiac arrests;
- decrease incidence of inappropriate resuscitation; and
- promote adoption and compliance with evidence-based practice

The regulatory bodies and collegiate societies discussed in this section all work very closely with the Department of Health. These organisations actively lobby the government and command significant influence in the development and implementation of healthcare policies. There is a general consensus amongst these organisations that evidence based practice, informed primarily by randomised controlled trial based studies and systematic reviews provide the best form of evidence to develop guidelines and healthcare policies. However, in practice, very few studies of this nature are carried out and the advisory bodies and committees of organisations such as NICE, NPSA and other related agencies use their expertise and experiences to inform practice. This has important implications for evaluations undertaken by NHS Trusts, which tend to use methods such as peer reviews, auditing and benchmarking against national standards. However, in reality, it would appear that decision makers trust the judgement and expertise of their fellow clinicians together with national guidelines.

### **4.3 Introduction to the study Trust**

The study Trust is one of the largest acute teaching NHS Trusts in England, catering for over one million people. At the time of data collection (from January 2012 to February 2014), it was working towards Foundation Trust licensing with Monitor. Attaining Foundation Trust status gives NHS Trusts independence from the Department of Health. The study Trust has two hospital sites, a district general hospital providing services to the local population and a tertiary centre, which provides complex specialist services to the local population and regional commissioners. The study Trust is the principal teaching

hospital for a local Medical School, with whom they collaborate on medical education and clinical research. It is one of the busiest NHS Trusts in the UK. In 2013/14, 574,242 people were seen in outpatient departments, 176,485 in the Accident and Emergency department 142,389 inpatients and day cases, 5,995 babies delivered and 41,157 operations were performed in its theatre departments. During the same period, the study Trust invested £200, 000 to upgrade equipment in theatre departments; £1 million to refurbish the Cancer Centre and introduced new innovative treatments such as the Intensity Modulated Radiotherapy. In September 2013, the study Trust launched its “Getting Emergency Care Right” campaign, which reportedly resulted in 7, 000 more patients being seen, receiving their treatment and being discharged within four hours in comparison with the same period in the previous year.

The study Trust has an Information and Computer Technology (ICT) department, which is responsible for developing, implementing and maintaining electronic clinical systems, working with third party suppliers and specialised departments. In 2005, the study Trust’s ICT department developed an award winning clinical results reporting system. It is an electronic patient record system (EPRS), which is widely used by doctors and an increasing number of nurses to obtain patients’ clinical history, ordering tests, viewing blood results and radiology reports amongst other uses. The EPRS is accessible through the study Trust’s intranet on desktop computers at nurses’ workstations as well as computers on wheels in ward areas. The ICT Department also provides training and support to users, as well as working with the Clinical Governance, Audit and Effectiveness and Information Departments to produce information for internal decision makers and regulatory authorities. Like many NHS organisations, the study Trust uses various indicators to monitor its performance through peer reviews and audits against national standards such as the NICE Pathways. Peer reviews are commonly used to evaluate clinical pathways and

operational documents to ensure conformity with published standards such as NICE guidelines and reduce variability in the study Trusts' departments. Clinical audits are used to establish whether the structure, processes or outcomes of services being provided meet national standards. Where shortcomings are identified, action plans are developed and changes implemented at ward, departmental or Trust-wide levels, with provisions for continuous monitoring to ensure full compliance. Clinical audits are part of the study Trust's wider Clinical Governance framework, which is used to improve the quality of care provided and patient outcomes. The study Trust's Audit and Effectiveness department keeps a registry of internal audits and its facilitators provide expertise and support to clinicians and managers to undertake and report clinical audits systematically. Root cause analysis techniques are used to systematically investigate claims, complaints and patient safety incidents to identify their cause, learn from them and ensure that similar incidents do not occur elsewhere. The study Trust and wider regulatory structure discussed in this chapter applies to all case studies selected for this research.

#### **4.3.1 Selected clinical decision support systems**

The next three chapters will describe the selected CDSSs, the clinical environments where they were used and evaluations that were carried out in respect to these CDSSs. Relevant aspects of the regulatory systems described in this chapter will be discussed in each individual case.

## Chapter 5 T1 Risk Assessment Tool

This chapter looks at the evaluations that were carried out for an electronic risk assessment tool (T1) using the CDSS evaluations framework that was developed in the literature review. T1 was used to assess patients who were at risk of developing venous thromboembolism (VTE) in acute hospital settings. A brief overview of the wider venous thromboembolism landscape is given in the introduction, noting some key studies. A wider overview of the management of VTE across the NHS is given in Section 5.2, noting the development of related national guidelines that affect VTE management. The study Trust's overall management of VTE and T1's development and implementation are discussed in Section 5.3. The evaluations that were carried out for T1 are discussed in Section 5.4. The key aspects of these evaluations are discussed in Section 5.5, noting similarities and differences with those identified in the CDSS evaluation framework.

The participants who were interviewed for this study are identified as follows:

1. T1 Project Lead Consultant – T1LeadConsultant
2. T1 Project Lead Nurse – T1LeadNurse
3. T1 Project Implementation Nurse – T1ImplementationNurse
4. T1 Lead Developer - T1Developer
5. T1 Charge Nurse – T1ChargeNurse
6. T1 Ward Sister (Medicine) – T1WardSister
7. T1 Ward Manager – T1WardManager
8. T1 Senior House Officer – T1SHO
9. T1 Senior Registrar – T1Registrar
10. T1 Advanced Nurse Practitioner – T1ANP
11. T1 Ward Sister (Combined Specialties) – T1WardSister1

## 5.1 Introduction

VTE is a common condition in hospitalised patients. It relates to the formation of blood clots (thrombosis) in the deep veins of the legs causing deep vein thrombosis, which may travel up the blood stream into the lungs and cause pulmonary embolism. It causes complications that are sometimes fatal and places considerable resource burden on healthcare systems worldwide (Anderson and Spencer, 2003; Geerts et al., 2008; Geerts et al., 2001; Geerts et al., 2004).

### 5.1.1 Venous thromboembolism clinical interventions

Commonly used preventative VTE treatments (thromboprophylaxis) for hospital in-patients at risk include anti-embolic compression stockings and low molecular weight heparin such as Enoxaparin®. However, two major studies; the multi-centre CURVE (Kahn et al., 2007) study and the multi-national ENDORSE (Cohen et al., 2008) studies, concluded that most eligible high-risk patients did not receive appropriate thromboprophylaxis treatment despite strong evidence of its safety, efficacy and cost effectiveness. Consequently, Douketis and Moinuddin (2008) argued that VTE remained a major and often unrecognised cause of complications and death in hospitalised patients. Although there was some disagreement (see, for example, Vardi and Haran (2012)), the majority of published studies and influential collegiate groups such as the British Thoracic Society, British Committee for Standards in Haematology, the Royal College of Obstetricians and Gynaecologists, the Scottish Intercollegiate Guidelines Network and the American College of Chest Physicians cited overwhelming evidence in support of thromboprophylaxis unless contraindicated. They also developed specialty specific guidelines for VTE management. In February 2005, the UK House of Commons Health Select Committee enquiry estimated

that over a third of hospital in-patients in England were at risk of developing VTE, leading to 25 000 avoidable deaths annually and considerable long term disability (Health Select Committee, 2005). The enquiry identified significant inconsistencies in the application of existing VTE guidelines, particularly the administration of thromboprophylaxis. The Health Select Committee enquiry also concluded that there was lack of awareness of VTE as a clinical problem across the NHS.

## **5.2 Overview of VTE Management in the NHS**

Following the House of Commons report (Health Select Committee, 2005), collaborative work commenced involving the newly formed All-Party Parliamentary Thrombosis Group, the Department of Health, the National Institute for Health and Care Excellence (NICE) and the Royal Medical Colleges. They sought to develop a systematic approach to VTE management across the NHS. Their aim was to educate and spread awareness of VTE as a clinical problem and to highlight the effectiveness of thromboprophylaxis. Furthermore, all NHS Trusts were tasked with developing multidisciplinary Thrombosis Committees and specialist thrombosis teams to champion VTE and to implement protocols locally using the existing clinical governance frameworks and audit systems. Following the Health Select Committee enquiry, the Department of Health commissioned NICE to develop VTE guidelines for surgical patients. In addition, the Health Select Committee also suggested that national guidelines should be developed for adult surgical in-patients, covering risk assessment, effective thromboprophylaxis and patient counselling. They recommended that VTE demonstration sites should be set up to showcase NHS Trusts that had a proven track record for excellent VTE management, including education, audit and quality control systems. Moreover, the Healthcare Commission was tasked with inspecting VTE compliance in NHS hospitals to ensure that all recommendations had been put in place



and that guidelines were being followed. Around the same time, the Department of Health's Chief Medical Officer wrote to all NHS England Medical Directors announcing the appointment of a multi-stakeholder independent expert working group to develop and implement a national VTE treatment and prevention strategy. He also recommended that all NHS Trusts should implement existing guidance from collegiate groups while the NICE guidelines were being developed.

### **5.2.1 Guidelines for prevention and treatment of VTE**

This section outlines in detail the VTE England Programme Timeline for rolling out national VTE management programme (VTE England, 2012), and the implications of national events for NHS Trusts. The first NICE guidelines for the prevention and treatment of VTE were published in April 2007. These guidelines covered in-patients undergoing surgery and were primarily based on existing evidence from randomised controlled trials. NICE recommended effective risk assessment and use of both mechanical and pharmacological thromboprophylaxis where indicated. The NICE VTE Guideline Development Group also proposed high quality monitoring of VTE-related adverse events and emphasised the need to ensure effective translation of research into practice in order to curtail the various uncertainties and inconsistencies that had been highlighted in the House of Commons Health Select Committee report. The independent expert working group also published its report in April 2007. It recommended mandatory VTE risk assessment for all hospitalised patients. In a letter sent to all NHS England Medical Directors, the Chief Medical Officer supported the independent expert working group's recommendations and announced plans to implement a national, systems-based VTE prevention strategy across NHS Trusts in England. This strategy became known as the VTE Prevention England programme. It was headed by NHS England's Director of Patient Safety, working closely

with the All-Party Parliamentary Thrombosis Group, the Department of Health, the national VTE Board, NHS commissioners and Medical Directors from all NHS England Trusts.

In January 2010, NICE published revised VTE guidelines and quality standards aimed at systematically improving the structure, processes and outcomes of NHS Trusts' VTE management efforts. Unlike the 2007 guidelines, which only focused on surgical patients, the new NICE VTE guidelines required that all adult hospital in-patients should be risk assessed on admission, and whenever their condition changed during their stay in hospital and on discharge as specified in the revised Department of Health's VTE risk assessment template. In April 2010, the Academy of Medical Royal Colleges issued a statement in support of ongoing government initiatives for VTE prevention, the updated NICE guidelines for VTE and quality standards, the national VTE risk assessment template and the then newly proposed Commissioning for Quality and Innovation framework. The academy reiterated to its members and fellows the importance of effective VTE risk assessment in primary and secondary care, as well as appropriate documentation, prescription and administration of thromboprophylaxis where indicated. The Academy members were tasked with auditing their local systems in line with the Commissioning for Quality and Innovation goals. Additionally, healthcare professionals were encouraged to ensure that VTE prevention became an integral component of undergraduate and postgraduate medical teaching, and wider dissemination through e-learning modules and other promotional materials.

In February 2012, the updated American College of Chest Physicians guidelines were released. They focussed on improving risk stratification methods for medical patients and those undergoing surgery, as well as advice for long distance travel and the use of new

anti-thrombotic treatments. The Department of Health issued further detailed VTE guidelines in March 2013. The new guidance stressed that VTE compliance data should include all patients admitted to hospital who had been risk assessed using the specified criteria. Around the same time, the NICE pathways for VTE were published. The pathways brought together existing NICE VTE guidelines, quality standards and all published materials used to support the national VTE prevention programme. In line with previous All-Party Parliamentary Thrombosis Group recommendations, the national VTE Board issued a commissioning toolkit to help commissioners make appropriate decisions locally to support effective VTE prevention using the NICE Quality Standards and Outcomes Framework. Further guidelines were issued by NHS England through a 'frequently asked questions' press release for commissioners and the NICE Quality Standards for VTE prevention aimed at improving and speeding decision making processes and audit systems at local level. VTE risk assessment and root cause analysis data collection and reporting were now a central feature on NHS Trusts' quality agendas. Figure 5.1. shows the Department of Health's national paper based VTE risk assessment tool.

## RISK ASSESSMENT FOR VENOUS THROMBOEMBOLISM (VTE)

Mobility – all patients (tick one box)	Tick		Tick		Tick
Surgical patient		Medical patient expected to have ongoing reduced mobility relative to normal state		Medical patient NOT expected to have significantly reduced mobility relative to normal state	
Assess for thrombosis and bleeding risk below				Risk assessment now complete	
<b>Thrombosis risk</b>					
<b>Patient related</b>	Tick	<b>Admission related</b>			Tick
Active cancer or cancer treatment		Significantly reduced mobility for 3 days or more			
Age > 60		Hip or knee replacement			
Dehydration		Hip fracture			
Known thrombophilias		Total anaesthetic + surgical time > 90 minutes			
Obesity (BMI >30 kg/m <sup>2</sup> )		Surgery involving pelvis or lower limb with a total anaesthetic + surgical time > 60 minutes			
One or more significant medical comorbidities (eg heart disease; metabolic, endocrine or respiratory pathologies; acute infectious diseases; inflammatory conditions)		Acute surgical admission with inflammatory or intra-abdominal condition			
Personal history or first-degree relative with a history of VTE		Critical care admission			
Use of hormone replacement therapy		Surgery with significant reduction in mobility			
Use of oestrogen-containing contraceptive therapy					
Varicose veins with phlebitis					
Pregnancy or < 6 weeks post partum (see NICE guidance for specific risk factors)					
<b>Bleeding risk</b>					
<b>Patient related</b>	Tick	<b>Admission related</b>			Tick
Active bleeding		Neurosurgery, spinal surgery or eye surgery			
Acquired bleeding disorders (such as acute liver failure)		Other procedure with high bleeding risk			
Concurrent use of anticoagulants known to increase the risk of bleeding (such as warfarin with INR >2)		Lumbar puncture/epidural/spinal anaesthesia expected within the next 12 hours			
Acute stroke		Lumbar puncture/epidural/spinal anaesthesia within the previous 4 hours			
Thrombocytopenia (platelets < 75x10 <sup>9</sup> /l)					
Uncontrolled systolic hypertension (230/120 mmHg or higher)					
Untreated inherited bleeding disorders (such as haemophilia and von Willebrand's disease)					

Figure 5.1 Department of Health national VTE risk assessment template (The Department of Health, 2010)

### 5.2.2 Patient safety initiatives for VTE prevention

In September 2008, the Department of Health published a national VTE risk assessment template to be adapted by NHS Trusts and used for all patients on admission to hospital. This was followed by the influential Map of Medicine's VTE prevention pathway, which

presented step-by-step criteria for the various specialty work streams within NHS hospitals and incorporated a VTE e-learning package. Map of Medicine's pathways collated evidence from wider literature and guidance from various regulatory and professional bodies into step-by-step processes that could be implemented locally at Trust level. Around the same time, the King's Thrombosis Centre launched its website to provide a central resource for VTE prevention by sharing best practice across the NHS. The King's Thrombosis Centre was selected to lead the Exemplar Centre Network to showcase good practice in VTE management as part of the national VTE treatment and prevention programme. The Exemplar Centre Network recommended various innovative methods to improve VTE prevention such as electronic VTE risk assessments, web-based learning modules, advocating for the inclusion of VTE risk assessment as a key performance indicator for NHS Trusts' clinical scorecards and developing the VTE Link Nurse Network to champion VTE prevention at ward level. By the end of 2013, 26 NHS Trusts across England had been accepted into the VTE Exemplar Centre Network.

In January 2009, the World Health Organisation published a surgical patient safety checklist to be used for all patients admitted into theatre as part of its global initiatives to promote surgical safety. The checklist was adapted for NHS Trusts in England and Wales by the National Patient Safety Agency (NPSA) into a single page document to identify the patient and the procedure to be undertaken and to document every phase of the surgical process from anaesthesia, the actual surgical intervention itself and post-surgery management, including VTE risk assessment. The inclusion of VTE in the World Health Organisation checklist was important because it acknowledged VTE's significance as a patient safety issue on a global scale. Also in 2009, the independent expert working group, in collaboration with e-Learning for Healthcare launched the e-Learning Venous Thromboembolism (e-VTE) website. The website provided interactive VTE learning

resources using various media and case studies to support traditional learning methods. It was designed to be a reference source for clinicians to enable them to understand and effectively risk assess patients and administer appropriate thromboprophylaxis. Continuing with the patient safety theme, the 'National NHS VTE Leadership Summit' was held in June 2009 to discuss developments in VTE prevention since the publication of the Health Select Committee report in 2005. Another Leadership Summit was hosted by the Health Services Journal in March 2011, in collaboration with the All-Party Parliamentary Thrombosis Group and the Department of Health. This followed the recently published second quarter results of the national VTE audits, which showed some marginal improvements but fell far short of the set targets. (See Table 5.1.)

### **5.2.3 Dr Foster Intelligence hospital ratings**

In November 2010, Dr Foster Intelligence published its Good Hospital Guide (Dr Foster, 2011). The Good Hospital Guide rated NHS Trusts based on percentages of VTE risk assessments that had been carried out. The guide revealed that most NHS Trusts were not achieving the mandated 90% CQUIN goal. Indeed, some NHS Trusts were even failing to provide any information at all regarding VTE risk assessments undertaken. Around the same time, the Department of Health issued the first quarter results of VTE data submitted by 159 acute NHS service providers. Out of 2.7 million patients admitted, only 47% had been risk assessed on admission in July and although this increased to 57% by the end of the quarter, it was still significantly below the target of 90%. On the backdrop of these compliance results, the Operating Framework for the NHS in England 2011/2012 specified that VTE prevention would remain a national CQUIN goal. The NHS Outcomes Framework also included VTE incidence as a quality indicator. Subsequently, the NPSA published a 'How to Guide' to VTE risk assessment. The guide aimed to build on existing

resources developed by various stakeholders for the national VTE prevention programme. VTE Prevention England also launched an iPhone application based course designed to help pharmacists, nurses and junior doctors to understand VTE prevention, treatment and audit processes. A subsequent report by Dr Foster Intelligence argued for the introduction of clinical coding on admission to distinguish pre-existing conditions from hospital-acquired events such as VTE and thus improve accuracy of reported data.

#### **5.2.4 Incentives for VTE management in the NHS**

In 2011, various initiatives were implemented to support the national VTE prevention programme. The first of these initiatives was the NHS England operating framework for 2010/11 that specified VTE as a national Commissioning for Quality and Innovation (CQUIN) framework goal. The CQUIN payment framework was initially set up in 2008 to provide additional incentives to NHS providers for achieving set goals on specified quality standards agreed with commissioners. The VTE CQUIN goal required NHS Trusts to achieve at least 90% valid risk assessments for all adult in-patients. Performance against this target was to be measured monthly through a census of all patients admitted to hospital. Failure to meet the target was to be punished by withholding a percentage of payments against the Trust's contract value. Monthly CQUIN compliance reporting commenced in June 2010. Table 5.1 shows the results of VTE data that was collected by the Department of Health. The results show that most acute NHS Trusts were compliant with the minimum percentage required to secure CQUIN funding. However, it was not clear how these data were collected. The majority of NHS Trusts did not have CDSSs like T1. The project leaders argued that it was likely that these data were overstated because of the large numbers of patient cohorts who were excluded from VTE risk assessments.

Our Strategic Health Authority were producing reports where they were ranking high... It was interesting and slightly frustrating. You start seeing some Trust where you know their processes start reporting 98% [CQUIN compliance] success rates and when you talk to them, you realise they don't count day surgery and don't consider such patients as in-patients... That's actually not the spirit of what NICE want... It seems to me no one is looking hard at the figures.

T1LeadConsultant

Also, T1 project leaders noted that many NHS Trusts were only carrying out audits of VTE CQUIN compliance on discharge, rather than census counting of all patients admitted as required by the Department of Health. They also expressed frustration at the lack of enforcement by the government.

The UK Thromboprophylaxis Forum questions the accuracy of these figures [other NHS Trusts' compliance audits]... There is no question about our figures, we capture everyone [every patient] and we know exactly when they are risk assessed... Other [risk assessment by other NHS Trusts] risk assessments are done by coders after discharge... Some do random checks of say 100 notes [patient notes]... It should be every single patient, that's how the government wanted it... We do that but we are [ranked] number 14 in the country, and realistically, our methods are robust and we should be ranking higher... More should be done at national level. The technology is there... there is no reason why government cannot insist that everybody reports the same way.

T1LeadNurse



<b>Period</b>	<b>Total Admissions</b>	<b>Total VTE assessments performed</b>	<b>Percentage of admissions risk assessed for VTE</b>
2010/11 Quarter 2 (December 2010)	919,192	524,650	57.1%
2010-11 Quarter 3 (4 March 2011)	2,865,803	1,959,683	68.4%
2011-12 Quarter 1 (2 Sept. 2011)	3,057,451	2,572,284	84.1%
2011-12 Quarter 2 (2 Dec. 2011)	3,225,261	2,846,014	88.2%
2011-12 Quarter 3 (2 March 2012)	3,247,984	2,945,876	90.70%
2011-12 Quarter 4 (8 June 2012)	3,327,261	3,076,885	92.5%
2012-13 Quarter 1 (7 Sept. 2012)	3,250,115	3,035,060	93.4%
2012-13 Quarter 2 (7 Dec. 2012)	3,292,886	3,090,323	93.8%
2012-13 Quarter 3 (1 March 2013)	3,330,310	3,132,820	94.1%
2012-13 Quarter 4 (7 June 2013)	3,299,470	3,107,117	94.2%
2013-14 Quarter 1 (6 June 2014)	3,331,811	3,179,055	95.42%
2013-14 Quarter 2 (August 2013)	1,101,051	1,052,689	95.6%
2013-14 Quarter 3 (November 2013)	1,133,722	1,087,129	95.89%
2013-14 Quarter 3(2 May 14)	3,408,617	3,262,454	95.71%
2013-14 Quarter 4 (4 July 2014)	3,442,924	3,303,424	95.95%
2014/15 Quarter 1 (5 September 2014)	3,418,985	3,282,421	96.0%

Table 5.1 VTE risk assessment data collection for NHS acute providers (the Department of Health)

Another key initiative was the NHS Standard Contract for Acute Services Guidance for 2010/11 (Department of Health, 2010). The new contract required all NHS Trusts to undertake root cause analysis and to report all confirmed cases of hospital acquired thrombosis occurring within three months of an admission episode to commissioners. Root cause analyses for VTE also became a CQUIN goal, with separate auditing, reporting and payment arrangements between NHS providers and commissioners. The Department of Health issued additional guidance on how to use the CQUIN payment framework for both VTE prevention and root cause analyses, and again restated the government's

commitment to improve patient outcomes. In support of this initiative, the National Patient Safety Agency published tools that were aimed at helping NHS Trusts to utilise the root cause analysis methodology effectively in order to learn from incidents and improve patient safety. The NHS Standards Contract re-stated the government's commitment to the VTE prevention project. VTE risk management standards were included for the first time in mental health and learning disability services providing both acute and community care. In April 2011, the nursing and midwifery leaders from VTE Exemplar Centres launched the National Nursing and Midwifery Network website. The website supported the national VTE prevention programme by raising awareness and sharing best practice among the public and healthcare professionals. Around the same time, the All-Party Parliamentary Thrombosis Group published results from its fifth annual audit looking at NHS Trusts' VTE policies. The audit was based on a freedom of information request sent out to all Acute and Foundation NHS Trusts. The audit revealed that most NHS Trusts had written VTE policies and were compliant with most of the NICE VTE standards apart from reassessment of bleeding risk after 24 hours and providing patients with verbal and written information on discharge. It also highlighted the continuing improvements in VTE risk assessment data. However, many NHS Trusts reported challenges in data collection and sought guidance on cohort exclusions for low risk patients. They also called for outcomes-focussed reporting based on percentages of at risk patients receiving thromboprophylaxis instead of census reporting for all admissions as required by the CQUIN framework. NHS standard contract provisions, VTE audits and root cause analysis reporting were found to be poor. Some NHS Trusts suggested a national registry for VTE related death rates to improve the collection and reliability of data. Key recommendations of the report included the continuation of root cause analysis and VTE CQUIN goals, and outcomes based data collection.

The new Operating Framework for NHS England 2012/13 confirmed its commitment to VTE prevention by maintaining VTE CQUIN goals. It also included the Patient Safety Thermometer as an additional CQUIN goal to monitor improvements in VTE prevention. At the end of 2012, the All-Party Parliamentary Thrombosis Group issued a report analysing its work since inception in 2005. The report highlighted the challenges faced, successes and the need to maintain the gained momentum in national VTE prevention. Their annual survey results published in the same month highlighted various challenges, especially the fact that a quarter of NHS Trusts were still failing to achieve the 90% CQUIN targets for VTE risk assessments. They also noted the lack of clear outcome measures, commissioning difficulties and failure by NHS Trusts to provide patients and the public with adequate information. NHS Trusts, primary care providers and commissioners were urged to share information and utilise the NICE Quality Standards to improve patient outcomes. The All-Party Parliamentary Thrombosis Group restated its commitment to being a 'critical friend' to the national VTE prevention programme and helping to embed VTE prevention into routine NHS practice. The NHS Standard Contract for 2013/14 also restated that NHS Trusts should continue to perform VTE risk assessments and root cause analysis for hospital acquired thrombosis.

This section has outlined the key events that shaped the national VTE prevention programme and the roles of main stakeholders. In the next section, the study Trust's responses to national events and initiatives are discussed. The development of T1 will also be described and the way national level initiatives and guidelines affected decision making in the study Trust will be considered.

## 5.3 Management of VTE in the study Trust

The study Trust provided treatments for various VTE related conditions in accordance with patient presentations and specialty specific work streams. The treatment regimens were mapped into electronic clinical pathways that were available on the study Trust's Intranet. The pathways also explained the roles and responsibilities of the multi-disciplinary teams and the need to include patients and their carers in VTE management through education and counselling. Like many NHS Trusts at the time, the study Trust did not have a systematic approach to VTE management in place prior to publication of the Health Select Committee report in 2005. Following publication of the All-Party Parliamentary Thrombosis Group recommendations, the study Trust formed a multidisciplinary Thrombosis Committee in 2006 (co-chaired by T1LeadConsultant) that was tasked with improving the overall management of all VTE-related issues across the Trust, with particular emphasis on preventing hospital acquired thrombosis. The Thrombosis Committee was a permanent committee that met quarterly to review national VTE prevention initiatives and formulated local responses in the form of clinical pathways and audits. It was also responsible for updating operational documents and Trust-wide dissemination of new evidence. The Thrombosis Committee also oversaw the development and peer review of local VTE policies and protocols to ensure their concurrence with national guidelines, the latest research evidence and the views and recommendations of professional bodies. Figure 5.2 shows the study Trust's electronic VTE risk assessment operational processes and the roles and responsibilities of the different healthcare professional groups that were responsible for VTE management.



### 5.3.1 Roles and responsibilities of key stakeholders

VTE management in the study Trust relied on effective multidisciplinary working between senior managers and clinicians, doctors, nurses and pharmacists (see Figure 5.2). Doctors were primarily responsible for performing VTE risk assessments, prescribing thromboprophylaxis and routinely reviewing and leading responses to changes in the patient's condition while hospitalised. The risks of developing VTE were weighed against those of bleeding by assessing the patient's current clinical problems, pre-existing conditions and checking kidney function and platelet count to determine their suitability for pharmacological treatment. In cases where there were no contraindications, a small dose of Enoxaparin® was prescribed for administration once per day to prevent VTE occurrence. Where a patient had already developed VTE, pharmacological treatment was based on assessment of possible contraindications, measurement of body weight and relevant blood tests. Patients who were unsuitable for pharmacological treatment were fitted with antiembolic compression stockings to help stop the pooling of blood in the legs. Nurses were responsible for administering thromboprophylaxis and were also required to alert doctors when patients had not been risk assessed; when prescribed doses might be inappropriate; and when a patient's condition changed. Additionally, nurses were responsible for providing patients with VTE advice on discharge, including teaching patients and their carers how to fit antiembolic compression stockings at home correctly. Doctors and nurses were also required to educate and counsel patients on how to avoid VTE occurrence during their stay in hospital and following discharge. Pharmacists were responsible for routinely checking patients' drug charts to ensure that prescribed medications were suitable and that there were no contraindications. Where anomalies were found, pharmacists advised doctors and nurses but could also make necessary amendments on drug kardexes to ensure patient safety. Most clinical specialties

undertook ward rounds daily, in which medical teams, senior nurses and pharmacists visiting each patient to: review patients' progress and response to treatments; make adjustments to prescribed medications where necessary; make referrals to other specialties where required; make plans for discharge where appropriate; and the like.

### **5.3.2 Paper based VTE risk assessment tool**

VTE prevention efforts in the study Trust were initially focussed on risk assessment and thromboprophylaxis for patients undergoing surgery in line with existing guidelines at the time (NICE, 2007). In 2008, the study Trust adapted the Department of Health's paper-based VTE risk assessment tool. The paper-based tool was widely promoted through medical grand rounds, Intranet newsletters, departmental multidisciplinary meetings, Trust mandatory training and medical teaching. However, T1LeadConsultant noted that there was significant resistance to the paper-based VTE tool, particularly from surgical specialties who believed that their patients would be exposed unnecessarily to the risk of bleeding and post-surgery complications related to pharmacological thromboprophylaxis.

It's not an immediate cause and effect, like if I do this, the patient is going to get better in hospital. It is all about deferred consequences... One of the major challenges was getting some groups of Consultants and junior staff to realise the clinical benefit [of VTE risk assessments] and also that there was no option. This is a nationally driven process with major implications if we don't do it. It was a lot of effort and there were a lot of - shall I say heated discussions - with some departments about the benefits or otherwise as perceived... Unfortunately, the House of Commons paper was unhelpful because the data they generated was quite frankly, ludicrous. They quote 25,000 deaths per year and if you extrapolate that, it would mean that we were having well over 100 deaths per year in this Trust. We certainly were not having that. I would dispute that figure and most people agree it is inaccurate. Yes it will be

thousands, but not 25,000, and unfortunately if you start with a figure which is ridiculous, people will not buy in, whereas if it had been more realistic, it would have been better.

T1LeadConsultant

T1ImplementationNurse also argued that the resistance in some specialist areas may have been fuelled by genuine concern for the specific patient groups as well as individual “agendas”.

There was a lot of resistance from the orthopaedics department, not only here [the study Trust], but countrywide. There was also resistance in neurology, understandably, due to the nature of their patient population and other surgical specialities... Obviously they have their own agenda and their own viewpoints and it depends whether the Consultant or other clinicians are on board.

T1ImplementationNurse

T1LeadConsultant also noted that most clinical departments did not use the paper-based tool and regarded it as additional and unnecessary paperwork. Efforts to audit the paper-based tool involved manually going through thousands of patients’ clinical case notes and drug charts to establish whether risk assessments had been completed and if appropriate thromboprophylaxis had been prescribed and administered to the patient in accordance with the NICE guidelines. Annual audits performed from 2006 to 2009 showed the paper-based tool’s lack of usage and acceptance in most of the clinical areas. There were also diverse interpretations of NICE guidelines across the study Trust. Additionally, there was lack of funding to develop and monitor a Trust-wide VTE management programme.

To be brutally frank it was an uphill struggle [the paper-based tool]... the best we achieved was 37% [compliance] with the paper tool... very difficult to actually audit



everyone, with 1 100 beds in the Trust and about 10 000 patients per month... about a dozen people helping and going out to look at rates [of compliance]... it wasn't being widely used and that was really disappointing despite maximal efforts from a number of people...

T1LeadConsultant

In 2009, the NPSA surgical checklist was added to the study Trust's surgery pathway for use alongside the existing paper-based VTE risk assessment tool. Subsequently, VTE risk assessment was included in mandatory training for those newly joining the study Trust and annually for all existing clinical employees. However, the VTE risk assessment session was only allocated 15 minutes of a day long training session that covered many other high priority issues such as infection control and thus its impact was reported to be low. The study Trust had also recently changed the mandatory training (including VTE training) to an eLearning portal in line with shifting national standards. The T1 project leaders expressed concern that this new format of VTE training was unsuitable for most Trust employees.

We have quite a good package [of VTE training] included in the mandatory training. Unfortunately, it was perceived that there was lots of mandatory training which could be done electronically, a move towards the national VTE training tool. The Trust are saying we have to use that as part of the electronic staff record... a national tool in order to tick a box... its about 3 hours long, way beyond what most people need, besides myself, and possibly T1LeadNurse, I would argue that pretty much no one else needs that level of detail. In fact, it does not even mention VTE risk assessment... I have gone back to say I think we need to re-think...

T1LeadConsultant

T1ANP also noted the effectiveness of the face to face VTE mandatory training.

The session [VTE mandatory training] has quite a lot of shock tactics in it about how important it is for the hospital to comply - not in a nasty way but it drills home quite effectively the importance of it - the importance of nurses' role within it... It is reasonably effective.

#### T1ANP

In February 2010, the Thrombosis Committee became aware of the study Trust's mandatory responsibility to collect and report VTE data through the CQUIN framework from April onwards. For this, the study Trust was required to risk assess at least 90% of all adult patients admitted. Non-compliance would result in a reduction of £750,000 in operational income from commissioners. Following meetings with the study Trust's Chief Medical Officer and Management Board, a multidisciplinary VTE project team (T1 project team) led by T1LeadConsultant was set up to develop the study Trust's strategy towards achieving the VTE CQUIN goals. The T1 project team included Thrombosis Committee members, senior clinicians who had prior experience working with the information and communication technology department (ICT), ICT project managers, nurse leaders and divisional business managers. In addition the Strategic Health Authority covering the study Trust's region set up a VTE Steering Group that was made up of senior clinicians and managers to provide support and advice on implementation of VTE prevention projects across the region.

The T1 project team agreed that an electronic risk assessment tool was the best way of achieving the CQUIN goals. The team had initially considered developing an electronic tool that would be embedded onto T3 (see Chapter 7). T3 was a handheld electronic track and trigger CDSS that was widely used by nurses in the study Trust. The T1 project team

decided however to build T1 on the study Trust's existing electronic clinical results reporting system platform because it was widely used by doctors, who would be responsible for performing VTE risk assessments and prescribing thromboprophylaxis. Another advantage of using this platform was that it was developed internally and the T1 project team expected that this would make it cheaper and faster to develop and make any changes easier in the future. Additionally, this platform would provide an audit trail using existing parameters such as clinical specialty, consultant team, ward and individual patient levels. It was also decided that T1 would be based on the existing Department of Health template and the newly published NICE guidelines [NICE, 2010]. T1LeadConsultant developed the VTE operational policy at the beginning of 2010. The policy was revised in 2012 in line with the national VTE prevention policy, the updated NICE guidelines and additional guidance from the Royal Colleges. It outlined the national VTE prevention strategy, its implementation within the study Trust, and the roles and responsibilities of key stakeholders with respect to VTE risk assessment and use of T1 (see Figure 5.2). The study Trust's Chief Medical Officer approved the plan and gained the support and funding from the Trust Board. T1 was made a priority project for the ICT department. Around the same time, the Department of Health extended the CQUIN reporting deadline to the end of June 2010. This allowed the T1 project team more time to develop and test the tool before Trust-wide rollout.

I am under no illusion if it wasn't for the CQUIN money I don't think this could have happened [successful T1 implementation] and it's absolutely pivotal to have the support of the executives and that really helped as it started the ball rolling...once they signed up to it, finance became less of an issue...

T1LeadConsultant

### 5.3.3 The Technology

T1 is an expert clinical decision support system (CDSS) that was used to assess the risk of patients developing VTE in acute hospital settings. It was designed to assist clinicians to systematically assess VTE risk at the point of care for hospitalised patients. T1 was based on NICE guidelines for the prevention and treatment of VTE. It was developed by an in-house multi-disciplinary team of clinicians, nurses, ICT project managers and technicians and senior Trust executives. T1 replaced the paper-based VTE risk assessment tool which had been in use for several years in the study Trust. However, the paper-based tool was not widely used and thus did not provide the study Trust with adequate audit data to meet the expected requirements. T1 was expected to improve clinical decision making and patient outcomes as well as collection and presentation of data in a format that would satisfy the requirements of regulatory authorities and ensure that the study Trust secured funding for meeting VTE CQUIN targets as set by the commissioners.

#### How T1 works

T1 automatically populated the patient's demographic data from the study Trust's existing clinical results reporting system using a patient specific NHS number or locally allocated hospital number. To initiate the VTE risk assessment, users were required to input the patient's clinical data by ticking relevant boxes as applicable. T1 then processed the patient's data and recommended appropriate actions depending on the level of risk presented. These actions included prescription of appropriate thromboprophylaxis on the paper drug charts based on the patient's weight, pre-existing risk factors and blood results. Additionally, T1 also calculated the appropriate dose of thromboprophylaxis based on the clinical information provided. The user was then prompted to confirm that they had prescribed the thromboprophylaxis dose as recommended. In cases where it was clinically justifiable, users could override the T1's recommendations and add an

explanation using a free text box provided, noting the reasons for deviation. At this point, the top part of the patient's electronic profile changed to amber colour to indicate that VTE risk assessment has been initiated. Within 24 hours of admission, users were required to reassess the patient's risk, particularly bleeding risk or any change in the patient's condition. Where the risk was deemed minimal, the risk assessment was completed and the top part of the patient's profile changed colour to green. Where the risk of bleeding or other adverse event became imminent, then a further risk assessment was required before completion and the profile colour code remained amber. In instances where VTE risk assessment had not been initiated within 24 hours, the colour code automatically changed to red, indicating that VTE risk assessment was overdue. The "traffic light system" (Green, Amber, and Red) was incorporated into T1 because it was being used on the existing risk assessments within the study Trust and clinical staff were already familiar with the sequence of the colour codes. The study Trust's Information Department electronically collected and submitted to the national database the total number of monthly admissions and valid VTE risk assessments that had been carried out, i.e., those initiated and completed within 24 hours of admission.

#### **Expected benefits of T1**

T1 was customised to suit existing work streams in the study Trust and to ensure compatibility with existing legacy systems. The customisation was primarily done to enable the collection and reporting of audit data as required by the commissioners and regulatory authorities. The project team expected T1 to be safe to use and also that it would improve work processes and the overall VTE management system. It was the first such system to be developed in the UK and went on to win regional and national awards for innovation.

### 5.3.4 Implementation of T1 in the study Trust

Following efficacy testing and clinical validation of T1 (see Section 4), online user training packages were developed for doctors. On completion of the training, users were assessed online before they could carry out VTE risk assessments on patients. VTE awareness campaigns detailing the CQUIN requirements and correct usage of T1 were published on the study Trust's Intranet and added to the agenda of multidisciplinary team meetings. T1 was implemented at the beginning of April 2010, four months earlier than the reporting deadline in order to give the project team time to monitor its usage and compliance against the set CQUIN targets and with a view to resolving any problems arising. T1LeadConsultant revealed that there was significant resistance immediately, particularly from senior clinicians in the surgical specialties. They believed that the VTE incidence published in the Health Select Committee report in 2005 were exaggerated and that T1 would deliver little benefit to their patients.

There are significant groups of people who don't regard hospital associated VTE as something that happens to their patients and think of this process as bureaucratic and of little clinical benefit. Part of the problem is that VTEs usually happen weeks after the admission episodes... we didn't have a mechanism to feedback to some of those consultants [before the Root cause Analysis process], as often they will not be reviewing these patients until weeks later...

T1LeadConsultant

Despite the opposition, initial Trust-wide compliance following T1's implementation was 70%, compared with the 50% forecast by the T1 project team and 35% previously achieved using the paper-based VTE risk assessment tool. The T1 project team attributed the unexpected success to T1's mandatory nature as well as the awareness campaign and

training that accompanied its introduction. However, the wording of the VTE CQUIN framework inferred that all patients, including those admitted for minor procedures and routine tests in outpatients' clinics, would be risk assessed. This included patients who would normally be considered as being at low risk of developing VTE. This inevitably led to disruptions in routine processes such as the day surgery unit and other patient groups who were considered to be at low risk of developing VTE. T1 ImplementationNurse noted the extent of operational disruptions following T1 implementation.

There wasn't time for a pilot. It [T1] had to go [live]. It did work but there were departments where it absolutely caused mayhem. The way it was worded meant that every admission had to be risk assessed, even minor procedures... For example, in the Surgical Day Unit, day surgery lists for minor procedures were held back. On the day the tool went live, most of the VTE project team were away and I had to deal with the various issues.

T1ImplementationNurse

In response to these challenges (around July 2010), the study Trust and others in the region approached their Strategic Health Authority to seek permission to rewrite local policies to exclude low risk patient groups without incurring penalties and losing VTE CQUIN payments. A subsequent meeting between the Strategic Health Authority and all regional Medical Directors led to an agreement for cohort exclusion for day surgery patients on the basis that they were at low risk of developing VTE. To improve compliance further, the T1 project team introduced a restriction that denied access to the patient's clinical profile on the clinical results reporting system in cases where the VTE risk assessment had not been completed within 12 hours of admission. In cases of emergencies or where patient safety was paramount, permission to defer a VTE risk assessment for a further two hours could be obtained and temporary access granted to

the user to view the clinical results reporting system in the meantime. This sanction was remarkably successful. Within two months of implementation, the clinical results reporting system access blocking and exemption of day surgery admissions had improved CQUIN compliance to the required minimum of 90%.

In some areas, particularly surgical specialties and obstetrics and gynaecology, provisions were made for senior nurses began to perform VTE risk assessments following online training and assessment. However, thromboprophylaxis prescription remained the responsibility of doctors. T1ANP questioned the effectiveness of the online training for senior nurses who were carrying out VTE risk assessments.

I don't think it [online VTE training for nurses] was effective. It was almost like a task. I remember it being quite hard to work through, not that it was a difficult assessment, just not very user friendly. Those online assessments are sometimes like tick box exercises. How much that actually adds to the assessment being complied with, I really don't know. I don't think it adds much knowledge as to why it's important. I don't think the training itself added much to my knowledge... I can't even remember what was in that training... it just doesn't stick really

T1ANP

T1ANP argued that face to face VTE risk assessment training for assessors could have been more effective.

Maybe if it was delivered in a group environment where you were being talked through it and you could ask questions, it stimulates discussion - those things could be so open to different interpretations – in a group setting, rather than being left to just doing it yourself.

T1ANP



Following recommendations by the VTE Exemplar Centre Network, in January 2011 the study Trust appointed a VTE Specialist Nurse (T1LeadNurse) to champion VTE prevention at ward level in collaboration with the Trust's VTE Link Nurse Network. The role of the VTE link nurses was to champion VTE prevention in their respective ward areas. This initiative was expected to increase VTE awareness and also improve overall VTE compliance across the study Trust. Additionally, T1LeadNurse was also responsible for performing VTE root cause analysis with support from T1LeadConsultant as well as coordinating the study Trust's mandatory training and the Link Nurse Network.

Although the study Trust's VTE risk assessment figures continued to improve, the T1 project team reported variations in how T1 was used in various clinical specialties across the Trust. For example some theatre departments requested that VTE risk assessments be completed before patients were transferred to theatre, although the VTE guidelines recommended post-operative risk assessment instead (depending on the post-operative risks presented). In some cases, Advanced Nurse Practitioners (ANPs) were asked to initiate the VTE risk assessments in pre-theatre outpatient clinics to ensure that CQUIN compliance was met. This initiative was also meant to minimise the workload for doctors on the day of surgery. However, most interviewees (including the T1 project team leaders) expressed concern that the nurses could not prescribe thromboprophylaxis and in some cases this resulted in patients not having their Exnoxparin© doses because doctors assumed that risk assessments had been completed. However, T1LeadConsultant noted that his role was to facilitate and support other clinicians to meet CQUIN compliance rather than dictate how they ran their departments.

There are lots of ideas, for example, to block patients from theatre until the risk assessment has been done... The actual requirement is for post-operative

thromboprophylaxis, not preoperative [so] from a NICE compliance and thromboprophylaxis point of view, it doesn't make a difference. ... I am happy to support them [departments' compliance ideas], but ultimately, they are the ones who best understand their patients. I can't tell them how to do things. I only give them the minimum requirements, in a facilitative role.

T1LeadConsultant

However, these different interpretations had little effect on the study Trust's monthly CQUIN compliance, which remained above 90%. VTE prevention remained a key feature of the study Trust's annual quality accounts since 2010, in line with NICE and the Department of Health's recommendations. By the end of 2014, T1 had not required any technical alterations apart from maintenance of its clinical aspects, particularly additional training, ensuring correct usage and addressing ad-hoc operational issues, which were largely undocumented but said to be time consuming by T1 project team leaders. At the end of 2012, T1LeadConsultant presented a business case to the Trust Board seeking additional funding for an administrator and another VTE Specialist Nurse to ensure adequate oversight of all VTE-related operational issues. However, by the end of 2014, the funding had not yet been approved despite the increasingly demanding workload.

## 5.4 T1 evaluations

A number of evaluations were performed during the development of T1 and following its implementation. These evaluations are shown in Table 5.2 in chronological order. Some of these evaluations, particularly those that were carried out during its development, focused solely on T1's technical and patient safety aspects. Others took a broader approach and covered the study Trust's wider VTE management system. The majority of evaluations were undertaken internally by the T1 project team using peer reviews, clinical

audits and root cause analyses. Informal evaluations were also undertaken by various stakeholders in the course of their routine work.

Evaluation method	Timing and duration of evaluation	Evaluators	Evaluation Purposes	Target audiences	Results	Uses of evaluations
Peer review of the Trust's revised VTE operational policies	Undertaken at the beginning of 2010 and in 2012. Not clear how long it lasted	Senior Haematologists	To ensure appropriate translation of NICE guidelines into practice	Thrombosis Committee, Trust Board, Commissioners and Regulators	Trust VTE policy implemented	Trust VTE policy used to inform VTE tool software algorithm development
Peer review of paper-based VTE algorithms	During the VTE tool's development. Duration reported to be "a few weeks"	Senior clinicians in the VTE Project Team	To ensure appropriate translation of the Trust's VTE policy	VTE Project Team, Thrombosis Committee and Trust Board	Paper-based algorithms approved and used to develop VTE software algorithms	Paper algorithms developed into software algorithms
Validation of software rules and validation	During the VTE tool's development. Duration reported to be "a few weeks"	ICT technicians and project managers	To assess efficacy of the tool's software	VTE project team, Trust board and commissioners	Appropriate changes made as part of the development process	VTE tool implemented Trust wide following validation
VTE tool usability and user acceptance testing	For a few weeks in May 2010	VTE Project Team	To check usability of tool interface and rules with real clinical users	VTE Project Team and clinical users	Interface found to be user friendly and rules appropriate	VTE tool implemented Trust wide
VTE CQUIN compliance audits	Performed monthly	VTE project team	To check compliance with VTE CQUIN targets	VTE Project team, Thrombosis Committee, Trust Board, commissioners and Regulators	Trust achieved and maintained VTE CQUIN targets	CQUIN financial incentives secured
Enoxaparin® and Doppler scanning	Performed one year after VTE tool implementation	Pharmacy and Haematology Departments	To assess the cost-effectiveness of Doppler scans, prophylactic and	Departmental business managers and clinical leaders	Doppler scanning and Enoxaparin® use found to be cost effective and in	Maintenance of existing VTE policy and practice

cost-effectiveness audit	over a two week period		treatment does of Enoxaparin®		line with national averages	
External audit of the Trust's VTE management system	Performed one year following implementation of the VTE tool over one week	PricewaterhouseCoopers auditors	To test the robustness of the Trust's VTE management systems	Trust Board, commissioners and the Strategic Health Authority	VTE management reported as effective in Trust's annual quality accounts	VTE management system maintained
Root cause analysis of hospital acquired VTE	Commenced one year following VTE tool implementation. Undertaken and reported quarterly	VTE TILeadConsultant and Lead Nurse	To establish causes of VTE for hospitalised patients	Trust Board, commissioners and regulatory authorities	VTE causes identified and appropriate remedial changes made in the affected areas and Trust-wide where necessary	Root cause analyses continue and have now become a separate CQUIN requirement
Enoxaparin prophylaxis and treatment audit	Performed three years after implementation of the VTE tool. Audit lasted one month.	Project Lead Nurse, VTE link nurses and Audit and Effectiveness Department facilitator	To audit Trust performance against NICE quality standards for VTE	Link Nurse Network, VTE Project Team and the Audit and Effectiveness Department	VTE documentation found to be suboptimal and inconsistent in most cases	Recommendations made for practice changes in affected areas and Trust-wide where necessary

Table 5.2 Evaluations undertaken reported by participants in chronological order

#### **5.4.1 Peer review of study Trust's VTE operational policies**

At the beginning of 2010, the study Trust's VTE operational policy was peer reviewed by a senior Haematologist and approved by the Chief Medical Officer in line with the Trust's policy. The peer review was repeated in 2012 following the newly published NICE guidelines. It was unclear how long the review process lasted on both occasions. The purposes of the peer reviews were to assess whether the study Trust's operational policies were compliant with NICE guidelines and other published clinical standards. The review process involved a Senior Haematologist reading through the study Trust's VTE operational policies and comparing them with the key aspects of the NICE VTE guidelines for the management of VTE. The policies were found to be compliant with NICE VTE guidelines and were used as a reference point for the development, implementation and use of T1. The peer reviews were also important because the VTE operational policies were used by healthcare professionals and managers as a reference document in their routine clinical work. Apart from the clinical members of the T1 project team and the Thrombosis Committee, most interview participants were unaware of this peer review process although the policies were available on the study Trust's intranet. The peer reviews looked at the operational policies' conformity with the underlying guidelines that supported T1, rather than an evaluation of T1 itself.

#### **5.4.2 Peer review of paper-based VTE rules and algorithms**

In April 2010, the paper-based T1 rules and algorithms were reviewed by clinicians who were involved in its development. The purpose of the peer review was to assess their compliance with the study Trust's VTE operational policy and NICE guidelines for VTE management. Additionally, the T1 project team wanted to assess whether the rules would enable them to appropriately capture the information required to fulfil CQUIN

requirements for VTE data collection. The Thrombosis Committee also wanted to ensure that all of the study Trust's work streams and appropriate prescriptions were adequately covered and any amendments made before T1 software algorithms were developed. According to T1Developer, the peer review involved checking the accuracy of the rules and algorithms against NICE guidelines. The rules and algorithms were found to adhere with NICE guidelines and were subsequently translated into software algorithms to support the development of T1. All T1 project team members interviewed noted that the validation of the rules was an important step to prove that T1 adhered to NICE guidance. Apart from members of the T1 project team and the Thrombosis Committee, none of the other interview participants were aware of this peer review.

#### **5.4.3 Validation of T1 software rules**

T1's software rules were validated by T1LeadConsultant from April to May 2010. The purpose of the validation process was to assess whether T1 software worked as intended in line with NICE VTE guidance. Furthermore, the T1 project team also wanted to ensure that the software rules covered all the potential clinical scenarios in the various clinical specialties across the study Trust. According to T1developer, the review process involved "iteratively going through all the rules line by line over several weeks to check that the tool was working as intended". T1's prescription aspects were also tested to assess their compliance with NICE guidelines and to ensure patient safety. To achieve this, the T1 software development team used simulations with synthetic and anonymised real patient data. All the test results were reviewed and "signed off" by T1LeadConsultant. T1 software rules were found to be "robust" and the results were reported to the Thrombosis Committee and the Chief Medical Officer. The validated electronic software rules were used to develop T1. Apart from members of the T1 project team and the Thrombosis

Committee, none of the interview participants were aware of these evaluations. However, there was an assumption amongst participants that these tests had been carried out.

#### **5.4.4 VTE tool usability and user acceptance testing**

According to T1Developer, the project team carried out usability and user acceptance testing with “actual clinical users over a few weeks” in May 2010. The purposes of these tests were to assess T1’s interface and user friendliness. Additionally, the T1 project team also sought to assess T1 audit trails and its effectiveness as a data collection tool to satisfy CQUIN reporting requirements. Clinical users were asked to describe their experiences of using T1. Their feedback was relayed to the software development team and necessary changes made where appropriate. T1Developer described this process as a continuous communication link whereby clinical users reviewed and tested every stage of T1’s development process and gave real time feedback to the software development team. This process was repeated over “a few weeks” until all the possible clinical scenarios and prescription options were tested. T1 was found to be user friendly and implemented as planned. However, most of the interview participants were unaware of these tests but assumed that they had been carried out as part of the T1 development process. Also the clinical users who were involved with the usability and acceptance testing were members of the T1 project team. These tests could have been expanded to include nurses and doctors in real clinical settings where T1 was intended to be used to gain feedback from users outside the T1 project team.

#### **5.4.5 VTE compliance audits**

From June 2010, the study Trust started performing monthly audits in some of the clinical departments to check VTE compliance as required by the VTE CQUIN framework. The purpose of these audits was to check whether patients that were admitted to the hospital



had been appropriately risk assessed as required under the NICE guidelines and the VTE CQUIN framework. To meet the CQUIN targets and secure funding from the commissioners, the study Trust had to achieve a minimum of 90% valid VTE risk assessments every month. The audits involved an electronic census of all patients admitted to the hospital. Audit reports were auto-generated and published monthly by the Information Department on the study Trust's intranet by department, and then collated across the Trust. They were then submitted monthly via the Unify2 database to the Department of Health. The CQUIN targets were achieved by September 2010 following measures such as exclusion of patient cohorts that were not considered to be at risk and blocking users' access to the clinical results reporting system where risk assessments had not been completed within the stipulated timeframes. The study Trust was one of a few NHS Trusts to achieve and maintain these targets up to end of 2014.

The VTE compliance audits were focused on achieving the CQUIN goals, rather than evaluating T1 itself. The majority of the participants interviewed showed awareness of the VTE compliance audits, particularly those relating to their own clinical areas. Most nurse managers and doctors argued that more could have been done to assess T1's effect on clinical practice and patient outcomes, rather than just counting the number of patients who were deemed to have had a valid VTE risk assessment. One senior doctor felt that VTE evaluations should be carried out by individual departments using existing clinical structures and the clinical governance framework, with overall checks by the T1 project team annually to ensure concordance with national guidelines.

Each department should have its own evaluation systems ... we are all responsible for our patients, and we do that actually. I attend monthly meetings for my ward and look at results for the ward. Ask how we are doing. If we are not doing alright, what are we doing about it? That in itself is an evaluation. About further evaluation, I do

not think it's necessary as long as individual evaluations are being done. Maybe there should be a team looking at it at the end of the year, to see if we are following the guidelines.

T1Registrar

Apart from the T1 project team members, most participants noted that they generally did not look at monthly VTE compliance figures but monitored their respective departments' performance through Quality Improvement Plans and wider clinical governance platforms. Two ward managers noted that ward nurses considered issues such as infection control more important than VTE risk assessments. T1WardSister noted that her ward consistently achieved the monthly VTE compliance targets. However, she argued that the VTE compliance figures had become 'mundane' and 'lost meaning' once targets were achieved.

For other important issues for nurses, such as infection control, real achievements can be seen visibly; the ward gets a certificate for reaching infection control targets every 3 months. Monthly figures [for VTE compliance] become mundane and lose meaning once the ward achieves 100% [VTE CQUIN compliance].

T1WardSister

T1WardManager argued that because VTE compliance figures were reported per department rather than individual wards, key contextual issues relating to individual wards were potentially being ignored or overlooked.

I appreciate that for certain reports it is practical but for a roll-out like this we [wards] should be kept apart because we did a lot of work on this and other things and when you put all the results together, it should be clear which individual components of the department are not up to speed as other issues may not be taken seriously . . .

#### **5.4.6 Enoxaparin® and Doppler scanning cost-effectiveness audits**

At the end of 2012, the study Trust's Pharmacy and Haematology Departments undertook joint audits looking at the cost effectiveness of prophylactic and treatment doses of Enoxaparin® and the appropriateness of Doppler scanning requests for suspected cases of deep vein thrombosis. The audits lasted two weeks. They were triggered by a surge in the volume of Enoxaparin® usage across the study Trust. Treatment doses of Enoxaparin® were recommended for patients with a suspected or confirmed VTE, while prophylactic doses were administered for VTE prevention. Enoxaparin® treatment doses were administered to at risk patients at a ratio of 1.5mg/kg of a patient's weight, and prophylactic doses were typically 40mg per dose and thus cheaper.

These audits did not address T1 itself but the services that were supported by its usage and the implications of inappropriate diagnosis and prescriptions. Due to staff turnover in these departments, there was no documentation available for these audits. However, T1LeadNurse explained that the Enoxaparin® audit involved looking at the appropriateness of prescriptions against the standards set out in the NICE guidelines. For the Doppler scanning audit, the evaluators looked at the numbers of scans that had been requested and also whether the results of the scans were positive or negative. The audits found that the majority of prescribed Enoxaparin® doses were appropriate and comparable with national averages and that most Doppler scans were requested appropriately. Findings from these audits were presented in a multidisciplinary departmental meeting (Haematology, Radiology and Pharmacy Departments). The existing Enoxaparin® and Doppler scanning pathways were maintained. Only T1LeadConsultant and T1LeadNurse were aware of this audit. T1LeadNurse noted that the

results of the audits attracted a lot of interest during the multi-disciplinary meetings, especially from the Consultants. She noted that they (Consultants) were primarily interested in the “numbers” because that was how they judged the success or failure of any intervention.

Some patients were not being scanned on time and pharmacists noted that Clexane® was being used increasingly ... I did an audit to see how many patients were admitted, how many doses were prescribed and were we diagnosing more patients? Were we doing more scans? The audit showed that the service was getting busier, not doing any more than other Trusts. There were concerns about length of stay waiting for scans. Clexane® usage went up by 200%. ... Doctors were really interested in the figures about patient waiting times. The audit probably needs repeating...

T1LeadNurse

#### **5.4.7 External CQUIN compliance audit**

Towards the end of 2012, the regional Strategic Health Authority tasked auditors from a leading accounting firm to audit the study Trust’s VTE compliance figures. According to T1LeadConsultant, the motivation for the audit was not disclosed. T1LeadConsultant noted that the Strategic Health Authority may have wanted to confirm the robustness of the study Trust’s compliance reports. Alternatively, he argued that they may have known that “our reporting systems could withstand such scrutiny” (i.e., the systematic audit) and could then prove the robustness of the positive results to the Department of Health and commissioners.

The SHA (Strategic Health Authority) did send auditors to look at our figures. I don’t know what the motivation was for that - whether they were being asked to justify our figures or whether they sent them because they knew we could justify our

figures. The auditors literally went through the finer detail of those 30 cases [using patient notes]. I am sure if the same was done for other Trusts, they would produce some interesting results. ... The problem of having an electronic system is we were then victims of our own figures; our figures can't lie. If other Trusts deliberately don't deal with day surgery, it makes them look like they are doing a better job than us and then you are not comparing like for like and that is really frustrating.

T1LeadConsultant

The external auditors examined thirty sets of patient case notes in detail and assessed them against the NICE VTE standards and the study Trust's VTE treatment pathways. However, the audit results were not fed back to the T1 project team but were reported positively in the study Trust's annual report and quality accounts. Only T1LeadConsultant was aware of this audit. He noted that the abolition of the Strategic Health Authority in April 2013 may have been the reason for the lack of feedback to the study Trust.

#### **5.4.8 Root cause analysis of hospital acquired VTE**

From the end of 2012, all NHS Trusts were required to complete a root cause analysis for every patient who developed a hospital acquired thrombosis. The root cause analyses were to be reported every three months to the Department of Health. Hospital acquired thrombosis was defined as a VTE occurring within 90 days of an admission. A separate CQUIN incentive was agreed with commissioners for the achievement of at least 90% minimum compliance, with the expectation that the target would be increased to 95% in the future. The purpose of the root cause analyses was to provide a systematic and evidence based method of finding out what factors or events led to patients suffering VTE while hospitalised and also to learn from these experiences and implement remedial actions plans. The NPSA's "basic steps of the root cause analysis" model was adopted by

the study Trust (see Figure 5.3). T1LeadConsultant and T1LeadNurse noted that root cause analyses had helped to improve VTE compliance and to maintain CQUIN targets because most clinicians and nurses were now aware that VTE data were being monitored. They also revealed that the root cause analyses had unearthed operational and patient safety issues in some poorly performing clinical areas. In such cases, the T1 project team involved clinicians, nurses and managers from the affected areas in the development and implementation of remedial action plans. This was done to encourage ownership and responsibility for those who were directly involved in delivering patient care to ensure that similar failures did not recur in the future. These failures were also reported to have presented opportunities for the T1 project team to engage directly with “willing” stakeholders who were keen to make changes in their respective clinical areas.

Some departments repeatedly get RCAs (root cause analyses). One ward has identified two nurses to champion VTE and trial thromboprophylaxis stockings on the ward. I met with the matron and the nurses and got good feedback from the meeting. ... They want to turn it around from a ward with issues to one that does VTE risk assessments very well. ... They may continue to have some problems due to the nature of their patients, but they would have done the best they could do. ... I like talking to nurses about real patient outcomes. It strikes home with nurses. The human side gets people on board. Every RCA is a chance to approach wards and it will mean more than just talking about CQUIN targets, i.e., this [the VTE] has happened ..., the guideline says this ..., ... wasn't followed, you can't do that on every ward. I really have to hold back, but you can't turn down an opportunity to teach, disseminate. I just get excited if anybody else is interested - any interest really.

T1LeadNurse

However, despite the additional CQUIN incentives for the study Trust for meeting the root cause analyses targets, no additional funding was allocated to the T1 project team for the additional work that was involved. T1LeadConsultant and T1LeadNurse took this additional responsibility and often worked in their spare time to meet the deadlines for root cause analyses submission. T1LeadConsultant described the process of performing root cause analyses as “more onerous and time-consuming” than the VTE compliance audits. The root cause analysis process often involved going through paper-based clinical notes as well as electronic sources and sometimes meeting with nurses and doctors to frame their investigations.

The problem at the moment is that the RCA process is a national CQUIN and is much more onerous to actually do so inevitably a lot of our time and focus is to make it happen. Because the risk assessment tool [T1] is hitting the targets, we haven’t had a chance to go back to areas where there are non-compliances and trying to straighten things out. We don’t have enough capacity to do everything, in fact, we are requesting additional support ... Because T1LeadNurse is just one individual, she spends a significant amount of her time doing admin duties, which is clearly inappropriate. We are producing all these figures but there is no point producing all these figures for the sake of it. It is not of any benefit to anyone.

T1LeadConsultant

Despite the challenges, T1LeadConsultant and T1LeadNurse noted that this process helped to spread awareness of VTE as a clinical problem especially to clinicians who were initially sceptical about VTE prevention initiatives. They argued that this awareness helped to improve patient outcomes, although these had not been directly measured. T1LeadConsultant and T1LeadNurse expressed disappointment that most NHS Trusts were not including incidental VTEs such as those identified from MRI scans, (which were

sometimes as high as 4-5 per week for the study Trust). They argued that these omissions resulted in lower rates of hospital acquired thrombosis for other NHS Trusts, while the study Trust ended up looking worse in the national ratings. T1LeadNurse argued that the government “could do more to ensure more accurate reporting”.

The root cause analysis CQUIN is labour intensive and we are looking to catch all hospital acquired VTEs. Other Trusts are not looking at incidental VTEs such as MRIs and yet it’s a requirement. ... Because we are looking harder for them, we may end up looking worse. This is where statistics can be made to look anything you like [by other Trusts]. ... It feels like we are shooting ourselves in the foot.

T1LeadNurse

Figure 5.3 shows the basic steps of carrying out a root cause analysis.

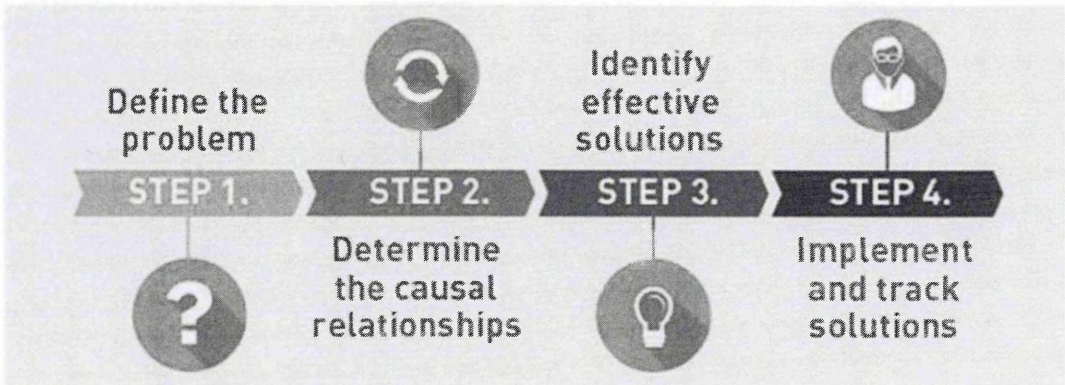


Figure 5.3 Basic steps of a Root Cause Analysis (adapted from the NPSA (2014))

**5.4.9 The Enoxaparin® prophylaxis and treatment doses’ audit**

In October 2012, T1LeadNurse and T1WardSister approached the study Trust’s Audit and Effectiveness Department for help with a “structured audit” to look at the use of anti-embolic compression stockings. The Audit and Effectiveness Department had previously received an National Patient Safety Agency alert in 2010 which highlighted that thromboprophylaxis was not being optimally used for high-risk patients, but had not been addressed. It was agreed that an audit looking at the interpretation and use of NICE



guidelines and the study Trust's wider VTE management system would address questions about the use of antiembolic stockings and concerns raised in the NPSA alert. T1LeadNurse and T1WardSister looked at 50 random sets of patient notes on four wards and the acute DVT clinic. The risk assessment, prescription and Enoxaparin® administration aspects were compared with the NICE VTE Prevention Quality Standard (NICE, 2010b). A summary of this standard is shown in Table 5.3.

The audits showed variations between the prescription, administration and documentation of VTE treatments. Generally, VTE documentation in clinical notes was found to be inadequate. There were also inconsistencies and incidences of poor application of T1. Patient safety issues arising from the audit were reported to respective clinical line managers and appropriate action plans were put in place. T1LeadNurse revealed that there were plans to repeat the audits and to develop other evaluation tools but resources were very limited.

Some [Doctors] were writing [Enoxaparin®] up but not ticking, 'and where recommended, Enoxaparin® not prescribed'. When contacted, some would say, 'it is just a paper/electronic exercise [the VTE risk assessment], but you will find the patient is getting the care'... but now on the system it says they should have Enoxaparin®, and it's a black mark on the Trust. Some doctors would write 'patient on Warfarin®', but if you tick that option [initially on T1] then the system won't recommend Enoxaparin® and it goes down on the system as not done.

T1LeadNurse

Quality Statement	Quality Measure	Implications for stakeholders
<p>All patients, on admission, receive an assessment of VTE and bleeding risk using the clinical risk assessment criteria described in the national tool.</p>	<p>Structure: Evidence that patients receive a risk assessment for VTE and bleeding that uses the clinical risk assessment criteria described in the national tool.</p> <p>Process: Proportion of patients assessed on admission for VTE and bleeding risk using the clinical risk assessment criteria described in the national tool.</p> <p>Numerator – the number of patients assessed on admission for VTE and bleeding risk using the clinical risk assessment criteria described in the national tool.</p> <p>Denominator – the number of inpatients and day cases admitted to hospital.</p>	<p>Service providers ensure patients, on admission, are assessed for risk of VTE and bleeding using the clinical risk assessment criteria described in the national tool.</p> <p>Healthcare professionals assess all patients for risk of VTE and bleeding on admission using the clinical risk assessment criteria described in the national tool.</p> <p>Commissioners ensure services assess all patients for risk of VTE and bleeding on admission using the clinical risk assessment criteria described in the national tool.</p> <p>Patients can expect to have their risk of VTE and bleeding assessed when admitted to hospital using the clinical risk assessment criteria described in the national tool.</p>
<p>Patients/carers are offered verbal and written information on VTE prevention as part of the admission process.</p>	<p>Structure: Evidence of written patient/carer information on VTE prevention being available to patients/carers as part of the admission process.</p> <p>Process: Proportion of patients/carers who are offered verbal and written information on VTE prevention as part of the admission process.</p> <p>Numerator – the number of patients/carers who are offered verbal and written information on VTE prevention as part of the admission process.</p> <p>Denominator – the number of inpatients and day cases admitted to hospital.</p>	<p>Service providers ensure verbal and written patient/carer information on VTE prevention is offered as part of the admission process.</p> <p>Healthcare professionals offer all patients/carers verbal and written information on VTE prevention as part of the admission process.</p> <p>Commissioners ensure services provide all patients/carers verbal and written information on VTE prevention as part of the admission process.</p> <p>Patients/carers can expect to be offered verbal and written information on VTE prevention as part of the admission process to hospital.</p>
<p>Patients provided with anti-embolism stockings have them</p>	<p>Structure:</p>	<p>Service providers ensure that services and protocols are in place to ensure that patients provided with anti-embolism stockings are fitted and</p>

<p><b>fitted and monitored in accordance with NICE guidance.</b></p>	<p>(a) Evidence of local arrangements to ensure patients provided with anti-embolism stockings have them fitted and monitored in accordance with NICE guidance.</p> <p>(b) Evidence of local arrangements to ensure that staff are trained in the use and monitoring of anti-embolism stockings.</p> <p>Process:</p> <p>(a) Proportion of patients with anti-embolism stockings fitted and monitored in accordance with NICE guidance.</p> <p>Numerator – the number of patients with anti-embolism stockings fitted and monitored in accordance with NICE guidance.</p> <p>Denominator – the number of patients provided with anti-embolism stockings.</p> <p>(b) Proportion of staff responsible for fitting and monitoring anti-embolism stockings who have received training on their use.</p> <p>Numerator – the number of staff who have received training in the use and monitoring of anti-embolism stockings.</p> <p>Denominator – the number of staff responsible for fitting and monitoring anti-embolism stockings.</p>	<p>monitored in accordance with NICE guidance by appropriately trained staff.</p> <p>Healthcare professionals follow local protocols and are trained in fitting and monitoring patients with anti-embolism stockings in accordance with NICE guidance.</p> <p>Commissioners ensure services are in place so that patients provided with anti-embolism stockings are fitted and monitored in accordance with NICE guidance.</p> <p>Patients who require anti-embolism stockings can expect to have them fitted and monitored in accordance with NICE guidance by appropriately trained staff.</p>
<p><b>Patients are re-assessed within 24 hours of admission for risk of VTE and bleeding.</b></p>	<p>Structure: Evidence of local arrangements to ensure patients admitted for more than 24 hours are re-assessed for risk of VTE and bleeding within 24 hours of admission.</p> <p>Process: Proportion of patients with a length of stay greater than 24 hours who are re-assessed within 24 hours of admission for risk of VTE and bleeding.</p> <p>Numerator – the number of patients who are re-assessed for risk of VTE and bleeding within 24 hours of admission.</p>	<p>Service providers ensure all patients with a length of stay greater than 24 hours are re-assessed within 24 hours of admission for risk of VTE and bleeding.</p> <p>Healthcare professionals re-assess patients within 24 hours of admission for risk of VTE and bleeding.</p> <p>Commissioners ensure services re-assess patients' risk of VTE and bleeding within 24 hours of admission.</p> <p>Patients can expect to have their risk of VTE and bleeding re-assessed within 24 hours of admission.</p>

	Denominator – the number of patients with a length of stay greater than 24 hours.	
Patients assessed to be at risk of VTE are offered VTE prophylaxis in accordance with NICE guidance.	<p>Structure: Evidence of local arrangements ensuring that the provision of VTE prophylaxis is in accordance with NICE guidance.</p> <p>Process: Proportion of patients assessed to be at increased risk of VTE who are offered VTE prophylaxis in accordance with NICE guidance.</p> <p>Numerator – the number of patients who are offered thromboprophylaxis in accordance with NICE guidance.</p> <p>Denominator – the number of patients assessed to be at increased risk of VTE.</p>	<p>Service providers ensure VTE prophylaxis is offered in accordance with NICE guidance having regard for the complications of thromboprophylaxis.</p> <p>Healthcare professionals offer VTE prophylaxis to all patients assessed as being at risk of VTE in accordance with NICE guidance having regard for the complications of thromboprophylaxis.</p> <p>Commissioners ensure services comply with NICE guidance on the provision of VTE prophylaxis having regard for the complications of thromboprophylaxis.</p> <p>Patients assessed as being at risk of VTE can expect to be offered VTE prophylaxis in accordance with NICE guidance having regard for the complications of thromboprophylaxis.</p>
Patients/carers are offered verbal and written information on VTE prevention as part of the discharge process.	<p>Structure: Evidence of written patient/carer information on VTE prevention being available to patients as part of the discharge process.</p> <p>Process: Proportion of patients/carers who receive verbal and written information on VTE prevention as part of the discharge process.</p> <p>Numerator – the number of patients/carers who receive verbal and written information on VTE prevention as part of the discharge process.</p> <p>Denominator – the number of inpatient and day case discharges.</p>	<p>Service providers ensure verbal and written patient/carer information on VTE prevention is available as part of the discharge process.</p> <p>Healthcare professionals offer all patients/carers verbal and written information on VTE prevention as part of the discharge process.</p> <p>Commissioners ensure services provide all patients/carers verbal and written information on VTE prevention as part of the discharge process.</p> <p>Patients/carers can expect to be offered verbal and written information on VTE prevention as part of their discharge plan.</p>

<p><b>Patients are offered extended (post hospital) VTE prophylaxis in accordance with NICE guidance.</b></p>	<p>Structure: Evidence of local arrangements to comply with NICE guidance on prescription of extended (post hospital) VTE prophylaxis.</p> <p>Process: Proportion of patients offered extended (post hospital) VTE prophylaxis in accordance with NICE guidance.</p> <p>Numerator: The number of patients offered VTE prophylaxis in accordance with NICE guidance.</p> <p>Denominator: The number of patients eligible for extended (post hospital) VTE prophylaxis.</p>	<p>Service providers ensure patients are offered extended (post hospital) VTE prophylaxis in accordance with NICE guidance.</p> <p>Healthcare professionals offer extended (post hospital) VTE prophylaxis to patients in accordance with NICE guidance.</p> <p>Commissioners ensure services comply with NICE guidance on the provision of extended (post hospital) VTE prophylaxis.</p> <p>Patients can expect to be offered extended (post hospital) VTE prophylaxis in accordance with NICE guidance.</p>
---	---	---

Table 5.3 NICE VTE Prevention Quality Standard (adapted from NICE)

## 5.5 Discussion

The findings from this case study will now be discussed in the context of the key factors of evaluations that were identified in the literature review as illustrated in Figure 2.2 (CDSS evaluation framework).

### 5.5.1 Purposes of evaluations

The main purposes of T1 evaluations were to assess its adherence with NICE guidelines, achievement of CQUIN targets, and its technical robustness, usability and patient safety aspects. Most of these evaluations were formative, and focused primarily on T1's ability to generate information for submission to the Department of Health, translation of NICE guidelines, technical and clinical accuracy and some usability issues. The target audiences of these evaluations were primarily regulatory bodies, commissioners, the Trust Board and senior clinicians. The clinical validation evaluations were fully funded by the study Trust, perhaps reflecting their importance and decision impact. None of the evaluations focused specifically on patient outcomes. However, T1LeadNurse argued that by showing that they were compliant with national guidelines and achieving CQUIN targets, they had also improved patient outcomes. Regardless, this was not necessarily the case as seen in the NPSA triggered audit which showed that a significant number of patients were not being treated in accordance with the NICE guidelines despite the study Trust having achieved and maintained the required CQUIN compliance. Also some patients who had been deemed appropriately risk assessed in the monthly audits were found to be over or undertreated. These revelations highlighted the shortcomings of evaluations undertaken, especially the narrow focus of T1 evaluation purposes.

The purposes of these evaluations can be viewed as aimed at providing accountability to key stakeholders such as the government, commissioners and patients through

compliance with NICE guidelines and achievement of VTE CQUIN targets. The evaluations also provided justification to the study Trust's decision makers for investing in the development and implementation of T1. Indeed a senior clinician who was also a member of the Thrombosis Committee argued that T1 was "one of the most successful projects...", and noted that "it achieved everything it set out to achieve..." This view was also shared by senior ICT managers and T1 developers, who considered T1 to be the most successful IT project they had implemented in the study Trust. T1Developer noted that the multi-disciplinary collaboration and the fact that T1 was "clinically, rather than ICT driven" had led to its success.

I would actually say it's [T1] one of the most enjoyable projects I have worked on. ...

It helped that T1LeadConsultant knew exactly what he wanted, he had clear specifications and knew the value of the project to the Trust as well as clinical practice. ... We didn't want it to be seen as just another cumbersome chore just to collect data as we wouldn't get any value out of it and nobody would use it.

T1Developer

### **5.5.2 Evaluation approaches and methods**

The main evaluation methods used were peer reviews and audits of the study Trust's performance against national standards before and after implementation. All the evaluations, apart from the root cause analyses, focused on single issues and did not address contributory factors that led to the reported outcomes. The limited evaluation methodologies may have contributed to the focus on single outcomes, which resulted in evaluations missing key contextual issues that may have provided better information to decision makers. The methodological limitations themselves may have contributed to the narrow focus of the evaluations undertaken.

No summative evaluations were carried out. This may have been because the main priority for the study Trust was to show that they were adhering to national initiatives and thus improving patient outcomes as set out in the NICE guidelines. T1LeadConsultant revealed that they had not evaluated “real benefits to patients” because of lack of time and funding. However, he argued that they had “largely cracked the nut” and would look at “real patient benefits” in the future.

We could always do more [evaluations], but it’s all about resources - its capability issues. I could spend more time and yes it would be better I’m sure. ... I feel that we have largely cracked the nut to a large extent. To improve further in a clinically meaningful way to patients would probably take a lot more effort. ... We are a fairly small team compared to others [the leading VTE Exemplar Centres]. Of course, they will be doing a lot better than us. If you ask how much, I would say the difference is very small.

T1LeadConsultant

The T1 project team had considered performing a “before and after study” to compare pre-T1 implementation audits with CQUIN audits to establish the extent of improvements. However, this project was abandoned because of lack of funding. Additionally, T1LeadConsultant noted that two other NHS Trusts had published similar studies around the same time which had shown improvements in the performance of VTE risk assessments of 20%. He did not feel that the study Trust could make any further contributions to the discourse. Although T1LeadConsultant accepted that there were differences between NHS Trusts, he believed that the findings from the two Trusts would be comparable to the study Trust. However, although the other Trusts may have been comparable in some ways, they had not developed and implemented electronic VTE risk assessment tools that were similar to T1. The study Trust may have missed an opportunity



to evaluate the effects of T1 for its own internal purposes. In that way, they could have learnt and shared knowledge with other NHS Trusts who may have been considering developing or implementing similar electronic tools. T1LeadConsultant noted that the haematology department lacked the necessary resources to carry out further evaluations and effectively utilise the data generated.

I regret that I've got data just sitting, pre and post T1. ... I could have checked to see if VTE rates had changed. ... I would have loved to publish that data. We have data from 2006, before any trust-wide policy, around 2008 with the VTE paper tool but I haven't had time to look at the raw data; it's quite labour intensive, all DVTs, all CT scans. To be honest, in the last year, there have been one or two other Trusts [who have published] ... so now not sure I am going to be able to add anything new to the discourse. You would hope that we are at similar levels [as the two other Trusts]. ... It would be very interesting but I am not sure I would add much.

T1LeadConsultant

### **5.5.3 Contextual and organisational issues related to evaluation**

T1 was implemented Trust-wide, without a pilot because the project team was working under tight deadlines. It is possible that a pilot or phased implementation could have allowed the project team time to consider T1's wider implications on staff and clinical workflow across the organisation. This could have possibly widened the scope of evaluations, duration and context while resources were still available to the T1 project team. All interview participants noted that there were many operational issues and resource shortages across the study Trust. They argued that these issues affected the uptake of T1 and how it was used and embedded into routine clinical workflow. However, although the T1 project team showed awareness of these challenges, none of these issues

were considered in any of the evaluations that were carried out. T1LeadConsultant argued that their primary focus was to ensure that the evaluations showed that the study Trust was compliant with NICE guidelines and to ensure that they provided enough information to secure VTE CQUIN funding. T1LeadConsultant also further noted that the challenges regarding “buy in” by intended users and resistance to T1 were expected with any change project.

There was resistance ... Very predictable, again, junior doctors are not the ones who will necessarily be seeing these patients when they develop VTEs. There is innate resistance to paperwork, especially when it is centrally generated. ... We did consider, do we get to a point where everyone should be on Clexane® unless there is a contra-indication? Sooner or later, they [patients] will come to harm by Clexane® undoubtedly. ... Unfortunately bitter experiences have shown us that if you don't have these kinds of systems, no one thinks about it at all. They will probably just write it [prescribe thromboprophylaxis] up and realise later that they shouldn't have written it up. ... trying to get people to stand back a minute ...

T1LeadConsultant

T1LeadConsultant also noted that compliance was improving because doctors were aware that there was an electronic audit trail.

Surprising thing is that, not all, but a significant number [of doctors], as soon as they know that it's their name, details, with time and date, they take it a bit more seriously. But of course, some will just do a shoddy job, or it may be they get caught up especially if there is an adverse event ... But as soon as it's their name, they know that there are potential implications ... As long as they are just scribbling their name, or signature which is illegible with no contact details, a lot of people don't worry too much.

Although T1 was reported to have improved the collection of VTE risk assessment data, its implementation did not necessarily result in improved clinical decision making. CQUIN targets were only a census of patients that were deemed to have been appropriately risk assessed. Apart from the financial benefits for the study Trust, achieving these targets did not necessary imply that patients had received adequate care. The audits and peer reviews did not address key issues such as T1's interoperability with existing legacy systems, compatibility with clinical workflows, effects on professional and multi-disciplinary relationships, the organisational changes instigated by T1 and the potential and actual disruptiveness to the organisation and the interdependencies between all these aspects. No comparisons were made with data from the years prior to T1 implementation so it was not clear how much impact T1 had on patient and organisational outcomes. Operational issues such as staff shortages and turnover, pressures on the hospital's bed capacity, increasing demand for services and constant top down changes within the study Trust also appeared to have contributed to low prioritisation of T1 in ward areas. Although members of the T1 project team spent several months briefing all clinical departments regarding how and why T1 was developed, its significance to the study Trust and training and assessing users throughout the implementation, there was lack of awareness of such processes amongst most interview participants. This may suggest lack of feedback from line managers to frontline staff, particularly ward-based nurses. It may also be a result of high turnover of junior doctors who moved to different departments within the study Trust and between hospitals in the region as part of their rotational training. Although adequate funding was made available for the VTE project at the development and implementation stages, no further funds were allocated for T1 evaluation following implementation. Furthermore, the complete lack of funding for the

VTE root cause analysis process further depleted already limited resources that were available to the T1 project team. This may be a reflection of the study Trust's lack of commitment to T1 evaluations beyond the immediate regulatory requirements. Additionally, failure to pass on CQUIN incentives for the root cause analysis process may have been a disincentive for managers and clinicians to carry out any evaluation activities beyond mandatory requirements that were specified in the CQUIN goals.

#### **5.5.4 Barriers to evaluation**

Although T1's compulsory nature was lauded for improving compliance, it may also have contributed to resistance by specialties such as Accident and Emergency and Surgical departments. T1ProjectNurse revealed that some clinicians only did the "bare minimum" that was required to "tick the box", while some deferred the risk assessment or gave inadequate reasons for not completing the risk assessment within the specified time. This resistance was reported to have been influenced by lack of acceptance of the VTE project itself and some key aspects of its evaluations, particularly the monthly VTE audits.

Understanding the clinical importance, whether using paper-based tool or T1, is an ongoing battle, particularly with nurses. Doctors do it because they are required to do it, but I do sometimes wonder whether they just pay lip service. Do they just do enough to change it [T1 portal] from red to amber?

T1ProjectNurse

The majority of the interview participants felt that evaluations such as the CQUIN audits were target driven. T1Registrar argued that these audits were "bureaucratic diktat" which did not benefit patients in any way and only served the needs of policy makers and the study Trust's Executive Board.

That's where politics comes in. I wouldn't want to do it [VTE risk assessments] just to keep somebody's figures high and to meet a target [and] to be politically right in this situation. I look at it from the patient perspective, not for the purpose of keeping managers happy. We have to understand that what we are doing is right, and that it is worth doing, but if it's purely chasing a target, I'm not sure that is something I subscribe to. I will do it because I feel as a clinician, that it is necessary, to prevent this. And there should be some more discipline if people don't do it voluntarily ... it has to be done with a little bit of blocking tactics [i.e. access blocking]. That is not something I prescribe to.

T1Registrar

Additionally the involvement of various government departments and arms-length organisations further fuelled these speculations. It could be argued that the barriers to T1's evaluation may have mitigated the purposes of evaluation, which in turn limited the methods employed and the scope of evaluations, duration and context undertaken. These barriers may also have limited the benefits of evaluations undertaken, especially the production of information that was useful for decision making rather than producing patient census figures for the Department of Health. The T1 project team argued that despite the shortcomings of their evaluation methods, by proving compliance with NICE guidelines and achieving CQUIN targets, they delivered their professional and organisational duty of care and in effect improved patient outcomes. This was indeed also the view expressed in the study Trust's annual reports and quality accounts since 2010.

### **5.5.6 Overview of T1 evaluations and missed opportunities**

There were many limitations of evaluations which resulted in various missed opportunities. The majority of evaluations that were adequately funded by the study Trust were formative. They were carried out during T1's development and their purposes were

to establish the functionality, safety and accuracy of T1. By failing to adequately invest in post implementation evaluations, the study Trust may have missed key information about the effects of T1 in clinical areas. Also by failing to undertake summative evaluations, the study Trust missed the opportunity to understand the wider implications of T1. Such learning could have informed both the VTE project itself and related future projects, as well as the wider healthcare sector. Table 5.4 summarises the opportunities that were missed by the T1 evaluations.

The evaluations undertaken were limited both in scope and methodology. They also failed to comprehensively address the key aspects identified in the literature review. The purposes of evaluation were primarily focused on T1's compliance with NICE guidelines and achievement of the VTE CQUIN targets. The methods of evaluation were limited to peer reviews and audits which were commonly used for service evaluation in the NHS. Although randomised controlled trials and other experimental methods were widely believed to be the "gold standard" for CDSS evaluation, none of the methods were used or even considered for evaluating T1. There was lack of funding for evaluations that were focused on patient outcomes, effects of T1 on users and its interoperability and compatibility with existing systems and wider clinical workflows. There were many barriers before T1 was even developed, during and following implementation. To overcome these barriers, the project team focused on benefits of evaluations, which in turn directed their purposes of evaluation and the evaluation methods employed. These decisions appear to have been made above the organisational context to satisfy the study Trust's regulatory obligations.

Key evaluation factors based on the CDSS evaluation framework	T1 evaluations carried out	Missed opportunities
Purposes of evaluation should cover a range of issues, rather than focus on single issues	The primary focus of evaluations was to collect and submit audit data to meet targets and secure funding from commissioners.	T1's effects on wider issues affecting patients, users and the Trust were not assessed
Various approaches and methods of evaluation should be applied throughout the CDSS's lifecycle	Primarily audits which were based on NICE and VTE CQUIN requirements, clinical validations, informal evaluations, and professional judgments	Qualitative studies that address issues that are considered important by key stakeholders across the Trust could have been used in addition
Contextual and organisational issues should be considered in any evaluation activity	Contextual and organisational issues only considered in informal evaluations	Formal summative evaluations could have been carried out to assess T1's wider effects and user satisfaction
Benefits of evaluation should be made clear to all stakeholders	Efforts made to spread awareness of the need for evaluations, expected patient benefits and potential financial losses to the study Trust	More emphasis could have been placed on patient and organisational outcomes rather than the potential financial losses, which were not considered a priority by most stakeholders
Barriers to evaluation should be mitigated by widening the purposes, approaches and methods, taking into consideration contextual and organisational issues as well as highlighting the benefits of evaluation	A range of barriers identified and addressed informally to improve buy in. Punishments such as blocking access to electronic records were used to improve VTE compliance audits	A wider range of methods could have been used to support formal evaluations

Table 5.4 Opportunities that were missed by T1 evaluations

## Chapter 6 T2 risk assessment tool

This chapter looks at the evaluations that were carried out for T2 using the CDSS evaluation framework that was developed in the literature review. T2 is an expert electronic clinical decision support system (CDSS) that was developed to help nurses monitor patients with stable prostate cancer following interventions such as surgery and radiotherapy. Sections 6.1 and 6.2 provide a broad overview of prostate cancer, its treatment and the management of patients with stable prostate cancer. Section 6.3 introduces the technology (T2) and sets out the context of the case study and describes how T2 was implemented in the study Trust. The evaluations undertaken pre and post adoption are described in Section 5 and in Section 6, these evaluations are examined in the context of the key factors of CDSS evaluation that were identified in the literature review (CDSS evaluation framework). The participants who were interviewed for this case study are identified throughout the chapter as follows:

1. Project Lead Consultant (Uro-oncologist/Surgeon) at the study Trust – T2LeadConsultant
2. Software Developer at the developing Trust – T2Developer
3. Specialist uro-oncology Nurse 1 at the study Trust – T2Nurse1
4. Specialist uro-oncology Nurse 2 at the study trust – T2Nurse2

Additional information from key actors in T2 promotional videos that are available in the public domain is also included. They will be identified as follows:



5. Lead Consultant (Uro-oncologist/Surgeon) at the developing Trust -  
T2DevelopingConsultant
6. Lead Nurse at the developing Trust – T2ProjectNurse

## **6.1 Introduction**

Prostate cancer is the commonest type of cancer affecting men in Europe. In England, Wales and Northern Ireland, over 36 000 new cases are diagnosed annually and around 9 000 deaths were reported (NICE, 2014). It is more prevalent in men aged over 65 years, particularly black men (National Cancer Intelligence Network, 2009). Hsing et al. (2000) noted a substantial increase in prostate cancer prevalence across the world since the late 1980s. Much of this increase was attributed to incidental diagnosis following transurethral resection of the prostate (TURP) surgery and the uptake of the prostate-specific antigen (PSA) screening test (Brewster et al., 2000; Bray et al., 2010). Despite the increase in its prevalence, Cancer Research UK (2014) reported that prostate cancer death rates have declined over the past two decades. However, they noted that the actual numbers of deaths from prostate cancer had risen due to the increase in the number of men surviving into old age. Some have attributed the decline in death rates to the wide uptake of PSA testing (Gilliland et al., 1994; Moore et al., 2009; Quinn and Babb, 2002). Others argued that there was little evidence to support this claim, noting that other factors such as improved cancer treatment regimens could explain these changes (Oliver et al., 2001; Coldman et al., 2003; Collin et al., 2008). A comparative study by Collin et al. (2008) looking at the incidence and deaths from prostate cancer between the UK and USA from 1975 to 2004 found no conclusive evidence that screening based on PSA testing reduced prostate cancer deaths

### **6.1.1 Prostate cancer interventions**

There are a number of interventions available for treating prostate cancer. These include active surveillance, radical surgery, radiotherapy, and brachytherapy and hormone therapies. However, there is wide variability in the natural history of prostate cancer. Many patients do not present any clinical problems, while others present with severe symptoms which require aggressive treatments. This variation leads to challenges in monitoring and measuring the clinical benefits and cost-effectiveness of interventions given to the various patient groups. Cancer Research UK (2014) noted that the main consequence of the increasing prevalence of prostate cancer was that there were now more men living with the disease who required long term monitoring. To address this problem, collegiate organisations such as the British Association of Urological Surgeons and the British Uro-oncology Group supported the uptake of NICE guidelines for prostate cancer (NICE, 2008; NICE, 2014) and also developed additional guidelines for their members. Likewise, cancer charities such as Cancer Research UK, Prostate Cancer UK, and arm's length organisations such as National Cancer Intelligence Network and the Prostate Cancer Research Centre worked closely with the Department of Health and NICE to improve the outcomes of patients with prostate cancer. Monitoring patients with prostate cancer involves the detection of spread or recurrence of disease, assessing symptoms, performing ongoing clinical examinations, PSA testing and managing psychosocial problems. T2 was designed to provide clinical decision support to help monitor this patient group.

## 6.2 Prostate cancer: the UK clinical context

Since the 1990s, the management of prostate cancer has been primarily focused on the timeliness of diagnosis. The timeline in Figure 6.1 outlines the key initiatives that were implemented nationally to support patients with stable prostate cancer. However, Moore et al. (2009) noted that despite the increase in PSA testing, the overall reduction in the number of deaths from localised prostate cancer had only marginally fallen. In the early 2000s, the focus shifted from diagnosis to the reduction of deaths from prostate cancer. The Cancer Intelligence Network (2009) analysed a cohort of 83, 701 men diagnosed with prostate cancer in England from 1999 to 2002 based on information from cancer registry records and the British Urological Surgeons' staging data. The study identified delays in screening and variations in the presentation of prostate cancers, especially where patients did not present any symptoms. They also reported that deaths from advanced prostate cancers were much worse than previously thought, and that relative survival significantly dropped after five years.

In England, Wales and Northern Ireland, the management of stable prostate cancer is primarily based on NICE guidelines and recommendations from the British Association of Urological Surgeons Section of Oncology, British Uro-oncology Group and the British Prostate Group. The first NICE guidelines for the diagnosis and treatment of prostate cancer were published in 2008 using evidence available from randomised controlled trials and systematic reviews (NICE, 2008). These guidelines emphasised the need for healthcare professionals to inform patients and their relatives adequately about prostate cancer diagnosis and the various treatment options available and offer advice on the effects of treatments on patients' quality of life, physical and psychological wellbeing. The NICE

guidelines also required NHS Trusts to ensure that appropriate referrals to specialist primary care services were made to address post-surgical complications such as incontinence and erectile dysfunction. The guidelines further recommended that patients who were unsuitable for primary care referral should be followed up in a hospital under the active surveillance program. Stable patients were to be put on a 'watchful waiting' regimen whereby treatment would be withheld until significant disease progression. NICE guidelines recommended that these patients should be followed up in primary care using locally agreed protocols. Those patients who had stable PSA levels for at least two years without significant complications following radical treatment were to be offered follow up in primary care by secure electronic communications. However, these patients could also be followed up in hospital-based clinics if they were participating in clinical trials. Regardless of whether these patients were seen in hospital or primary care, NICE recommended that patients should be offered direct access to the urological cancer multidisciplinary team and that their PSA levels should be checked annually. Patients who had stable PSA levels for at least two years with no significant complications following radical treatments and those on the watchful waiting regimen were said to have stable prostate cancer. This is the patient group that T2 was developed to manage.

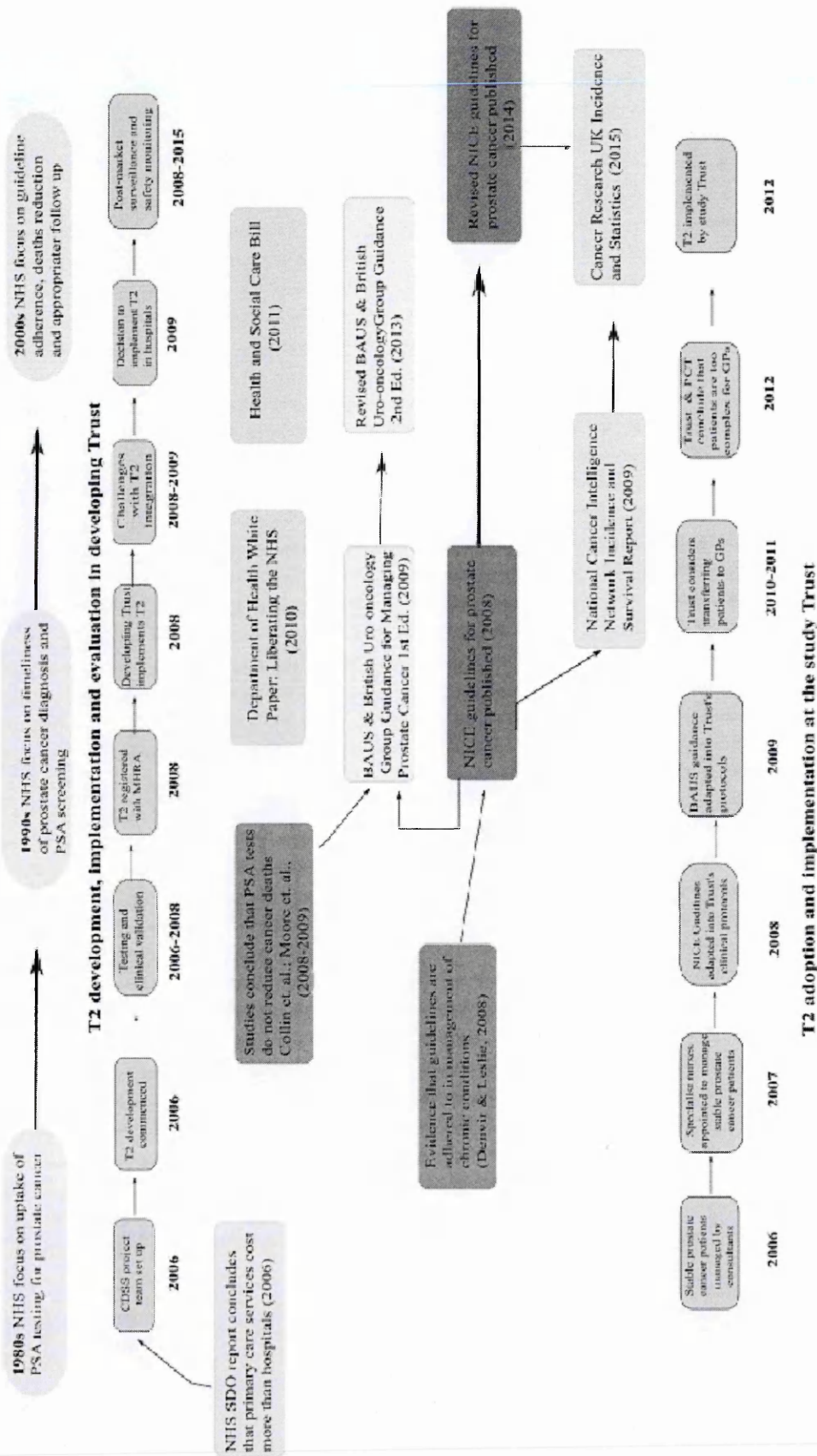


Figure 6.1 Timeline showing key prostate cancer initiatives and T2 adoption

Many of the issues covered in the NICE guidelines (NICE, 2008) primarily pertained to procedures that were undertaken in secondary care settings (NHS hospitals). They did not address the effects of service organisation and management of symptoms in primary care. Furthermore, there was lack of robust evidence to inform the NICE guideline development. Consequently most of the NICE recommendations were based instead on consensus opinion of the Guideline Development Group. Also in 2008, NICE published a costing tool and implementation pack to facilitate uptake of the new guidelines in the NHS. In November 2009, the British Uro-oncology Group, the British Association of Urological Surgeons Section of Oncology and the British Prostate Group published a multi-disciplinary guidance for managing prostate cancer (British Uro-oncology Group, 2009). The focus of the guidance was to promote integrated care to ensure consistency across the various stages of prostate cancer treatment. The guidance also sought to develop a structure that would facilitate auditing and peer review of prostate cancer services that were provided and to support the existing clinical governance framework in NHS Trusts. The British Uro-oncology Group guidance largely adopted NICE guidelines (NICE, 2008) and provided clarity on how they would be implemented throughout the patient's treatment journey. The British Uro-oncology Group guidance were revised in September 2013 (British Uro-oncology Group, 2013). The revised guidance emphasised that patients who were on watchful waiting regimens should be reviewed by the urological cancer multi-disciplinary team and sought to put in place structures for collaborative working. In January 2014, NICE published updated guidelines for diagnosis and treatment of prostate cancer (NICE, 2014). The revised NICE guidelines recommended wider prostate cancer management based on new evidence from observational studies, clinical and cost effectiveness studies, consensus surveys and randomised controlled trials. However, due

to the continued lack of evidence, new recommendations for the follow up of stable prostate cancer patients were mainly based on the experience and opinion of the NICE Guideline Development Group. The updated NICE guidelines included newly licensed treatments such as radiation-induced enteropathy. They also aimed to reduce variations and uncertainties that still existed in practice, especially the management of effects of radical treatments and investigations such as flexible sigmoidoscopy for patients with stable prostate cancer.

## **6.3 Management of stable prostate cancer in the study**

### **Trust**

The study Trust offers a wide range of services for the diagnosis, treatment, assessment and support for patients with prostate cancer. Emergency and acute admissions come through the Accident and Emergency department. They may be transferred to theatres for surgical procedures or urology ward for monitoring, further investigations and treatment. Elective patients are also admitted for planned procedures through the urology wards and the day surgery unit. GPs can refer patients who are suspected of having prostate cancer via fax or email, using a proforma that was recommended by NICE guidelines. According to the NICE guidelines all patients who are referred by GPs have to be seen within 14 days. A team of consultant uro-oncologists and specialist uro-oncology nurses performs minor procedures and monitors patients with stable prostate cancer and other urological problems in outpatients' clinics. The specialist uro-oncology nurses' role has expanded as the range of radical treatments offered has increased. Their additional responsibilities include health promotion, routine urological assessments, advising patients and their relatives regarding treatment options, symptom and side effects

management and supporting patients throughout their disease journey. Non-clinical support and advice is available from the MacMillan Cancer Information and Support Centre. There are also support groups that offer prostate cancer patients opportunities to meet with other people who have experienced prostate cancer. Additionally, service users are sign-posted to other community-based cancer services to access further support. Prior to T2 implementation, the stable prostate cancer outpatients' clinics were managed by two senior uro-oncology nurses with support from a team of consultant urologists. The study Trust's clinical pathways and protocols for the management of prostate cancer were based on locally developed clinical pathways, NICE guidelines (NICE, 2008), and recommendations from cancer charities and collegiate organisations. Appointments, referrals and follow up schedules were managed by the nurses with support from the consultant uro-oncologists' secretaries. The study Trust's electronic clinical results reporting system was used to check and update tests and results, make referrals and as a repository for clinic records and GP outcome letters. All clinic visits were also recorded manually by the uro-oncology nurses in the patients' paper-based medical notes which were also used by the uro-oncology multi-disciplinary team.

In 2011, the study Trust looked at the feasibility of transferring patients with stable prostate cancer for management by GPs in primary care as recommended in the NICE guidelines (NICE, 2008). The study Trust's Lead Uro-oncologist (T2LeadConsultant) worked with a senior General Practitioner who was the clinical lead for prostate cancer in primary care to assess the suitability of transferring this patient group from hospital to GPs. They found that most GPs lacked the expertise and capacity to effectively manage this patient group. They concluded that over 50% of the patients were unsuitable for management in primary care because they were too complex.



Only half of the patients were suitable for follow up by GPs but they [GPs] hadn't quite appreciated that. They thought all prostate cancer patients would be the same whereas it wasn't that. ... We then explained to the whole body of GPs to see if they wanted to manage these patients in primary care and they felt it was too complicated. They asked us to find a cheaper way of managing these patients.

T2LeadConsultant

Additionally, primary care services were undergoing significant structural changes at the time, which eventually resulted in the formation of Clinical Commissioning Groups to replace Primary Care Trusts. Consequently, local NHS Commissioners tasked the study Trust with developing a cost-effective strategy to continue following up these patients in hospital-based outpatient clinics. Figure 6.1 shows how the study Trust responded to various national initiatives that led to the implementation of T2.

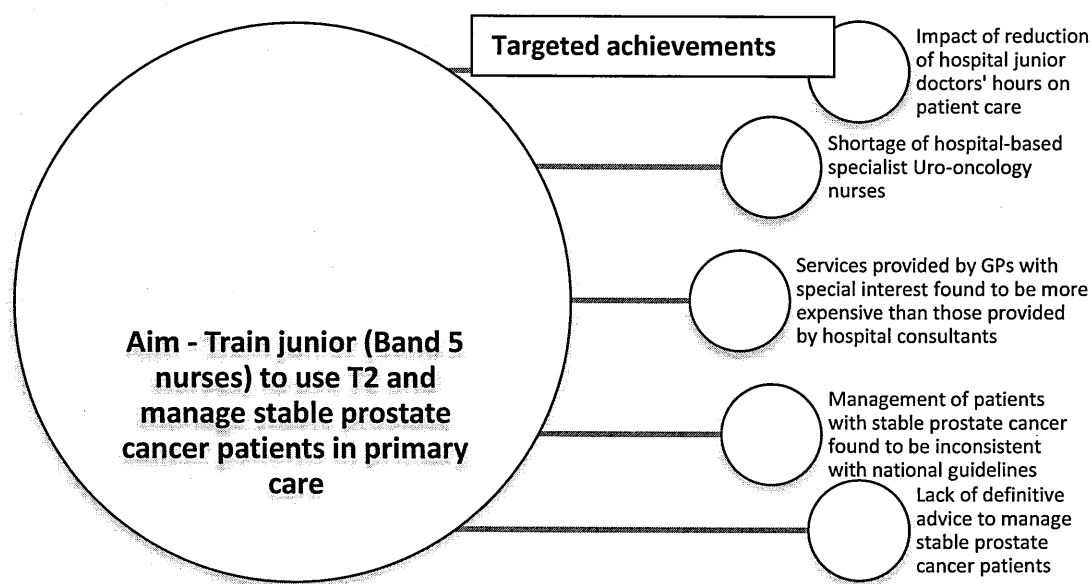


Figure 6.2 Problems that were expected to be resolved by implementing and using T2

### 6.3.1 The technology

T2 was developed over a two year period from 2006 to 2008 by a medical software development company, with support from a pharmaceutical company, a regional NHS Innovations Hub, a regional Development Agency and an NHS Trust (the developing Trust). The project partners provided resources to free up time for a multi-disciplinary team of software developers, specialist uro-oncological nurses, consultant uro-oncologists and NHS executives to support the T2 software development team. The developers focused on introducing various CDSSs to improve the management of chronic diseases in community settings and to reduce the costs of healthcare without affecting the quality of care. Prostate cancer was identified as one of the high priority disease domains. T2 was designed for use by Band 5 nurses with appropriate training to provide a nurse led service for stable prostate cancer management in local GP practices across the region. Band 5 nurses are junior nurses who are sometimes referred to as staff nurses.

T2 developers sought to promote 'Practice Based Commissioning' of NHS services, whereby GPs would take the lead in the commissioning of technologies to support their clinical work. The developing company's Medical Director (T2DevelopingCounsellant) was also a practising consultant uro-oncology surgeon in the developing Trust. He was involved in various healthcare IT projects and medical research projects that were linked with the local medical school and other regional partners. T2Developer argued that T2DevelopingCounsellant's diverse background and understanding of the NHS helped to support their concept of Practice Based Commissioning. These developments were against the backdrop of continuing increases in the aging population whose risk of developing prostate cancer was higher. Furthermore, there were also political factors

such the then proposed reforms that were eventually outlined in the government White Paper, *Liberating the NHS* (The Department of Health, 2010) and the *Health and Social Care* bill (The Department of Health, 2011), the rising cost of healthcare services and the shrinking public purse. T2 developers expected these issues to have significant effects on hospital services. The developers noted findings from a briefing paper published by the NHS Service Development Organisation which showed that services delivered by General Practitioners (GPs) with special interests (such as prostate cancer) were 50% more expensive than those delivered by hospital-based consultants (NHS SDO, 2006). The majority of the costs that were identified related to human capital costs. They also noted a research paper by Denvir and Leslie (2008), which showed that only 50% of heart failure patients were investigated and treated according to the national guidelines. The T2 developers argued that despite widely held assumptions that prostate cancer was easy to manage, it was a complex condition with over 30 different clinical scenarios. T2 was designed to counter these identified shortcomings by providing “definitive advice” for junior nurses to effectively manage patients with stable prostate cancer in primary care settings.

### **6.3.2 How T2 works**

T2 is a web-based expert clinical decision support system which was accessed via a secure online portal that is provided by the developers. Its software algorithms and system alert levels were based on NICE guidelines for the diagnosis and treatment of prostate cancer (NICE, 2008). It has built-in alerts, prompts and reminders, which are aimed at reducing the risk of missing key symptoms, ensuring appropriateness of referrals, reducing variability in practice and adhering to NICE guidelines. The software algorithms were

designed to allow amendments such as updates to guidelines and recommendations from clinical practice. The original design premise was that T2 developers would assist commissioners with the procurement process, provide the nurse to manage the T2 clinics as well as the required training and implementation. The initial patient clinic visit would be carried out face to face by the T2 nurse at the patient's local GP practice. For subsequent visits, the patient would then be offered a choice between a telephone (virtual) clinic or a visit in person at their local GP practice. During the initial clinic visit, the T2 nurse would input the patient's demographic and historical clinical data, ask quality of life related questions, assess clinician and patient reported symptoms and blood results, collate routine and specialist investigations and results, medications and treatments to date. T2 then processes the information and identifies trends that indicate disease recurrence or progression. It also suggests further tests where required and referral to a hospital-based consultant uro-oncologist where indicated. On conclusion of the visit, T2 generated a clinic outcome letter; one for the patient and another for the patient's GP. The outcome letter notes progress to date, suggests further tests and clinical management, follow up required and recommends booking of future tests where indicated. The outcome letter can be generated in multiple languages. It can be printed out and posted, emailed directly to the GP or even integrated into existing electronic patient record systems where possible. The key design criterion was to ensure that nurses spend very little time inputting data, but instead, focusing on the patient. The problems that were expected to be resolved by T2 are shown in Figure 6.2.

### 6.3.3 Expected benefits of T2

The overall aim of T2 was to enable the monitoring of patients with stable prostate cancer in primary care through nurse-led clinics. T2DevelopingConsultant noted that there were many “shared desires” amongst the various stakeholders that would be enabled by T2 implementation. Among these “shared desires” were cost reductions, improved patient outcomes and safety of care, timely referrals where indicated and data security. Using T2, patients were expected to receive the same quality of care regardless of whether they were followed up by hospital based uro-oncologists or primary care based T2 nurses. T2DevelopingConsultant argued that patients would receive localised “hospital quality care in the community”. The variability of care was expected to decline because adherence with NICE guidelines would improve through T2 use. The developers argued that T2 would provide a mechanism to identify problems and recommend appropriate action plans. Additionally, patients would also have a choice between telephone and face to face clinic appointment at their local GP practice. For the targeted T2 users (Band 5 nurses), the risk of missing key clinical indicators was expected to decline because T2 would allow the continuous collation of patient reported outcomes and clinical outcome measures, trigger key indicators and provide timely recommendations through in-built alerts and reminders. Availability of comprehensive clinical data was expected to enable decision makers to implement appropriate interventions where required. T2DevelopingConsultant noted that the inbuilt alerts and prompts would help to provide the required expertise and a “safety net” for nurses, thus reducing the risks associated with clinical decision making, while at the same time “upskilling” Band 5 nurses.

The developers expected the introduction of T2 into primary care to reduce the need for patients to attend hospital-based outpatient clinics. The number of patients re-referred

back to hospital inappropriately following interventions was also anticipated to decline. As a result, hospital-based consultant uro-oncologists and specialist uro-oncology nurses' time would be freed. T2DevelopingConsultant argued that consultant uro-oncologists would perform more operations, dedicate more time to trainee doctors and follow up only the most appropriate patients in outpatients' clinics. Specialist uro-oncology nurses would concentrate on more complex patients as well as performing other interventions in outpatients' clinics. These benefits were also expected to result in improved patient satisfaction, more efficient workload management and substantial savings for the hospitals.

T2DevelopingConsultant noted that the "T2 model" entailed that the developers would provide the "T2 nurse", delivered the required training and equipment, as well as assistance in developing the business case required to apply for funding. Furthermore, he argued that there would be significant tariff reductions and transparency through a "contracted in service model" (T2DevelopingConsultant). T2's inbuilt management tools would provide regular reports on patient outcomes. Commissioners were also expected to be reassured by T2's safety, which had been proven through clinical trials and following strict guidelines that were required for its registration with the MHRA as a software medical device and for CE marking. T2Developer noted that MHRA registration was required because T2 provided therapeutic effects to patients. MHRA registration was also important because it enabled the developers to show T2's robustness to potential commissioners and users. The developers argued that they had followed the same rigour as that observed in clinical trials for drugs and pharmaceutical products in evaluating T2. The developers also obtained clearance from the Care Quality Commission and level 2 clinical governance clearance as a third party supplier to the NHS. T2 also had support

from various cancer charities and related services and was mandated as a standard of care for NHS providers in the region where it was developed. The developers also argued that T2 was cheaper than existing alternatives, because Band 5 nurses would manage the stable prostate cancer clinics with T2 support instead of consultant uro-oncologists or other specialist nurses in hospitals. T2DevelopingConsultant noted that using T2, the NHS would save £500 million annually without affecting quality of patient care or in some cases even resulting in better care. T2Developer argued that T2 was “relatively easy” to commission using the developers’ templates which required completion with local data.

For many primary care settings, the developers envisaged working with various GP practices, with one site acting as a hub that would take referrals for appropriate patients from other practices. Commissioners would thus benefit from improved economies of scale and save money, with full reassurance of the effectiveness of treatments and interventions, improved quality of care, safety and improvement in outcomes. GPs were expected to improve their management of patients with stable prostate cancer, make substantial cost savings and reduce bureaucracy through the proposed ‘Practice Based Commissioning’ model. T2 was also expected to mitigate the effects of reorganisation of primary care services because it would introduce uniformity across various GP practices. GPs were expected to benefit from graphical presentation of clinical trends, cohort reporting and T2 customisation to local pathways and protocols. Additionally, T2 could be integrated with existing legacy systems to allow seamless sharing of clinical data between users and systems, avoid double entries and minimise human errors. Other expected benefits included alleviation of time pressures on GPs and mitigation of the lack of expertise to deal with this patient group in primary care. T2DevelopingConsultant argued that there were too few GPs with special interest in the management of prostate cancer

and that T2 would bridge the knowledge gap because it was “not too complex for the average GP” to use.

In their efforts to rollout T2 in primary care, the developers found that patient data were spread over various servers and often GPs used different systems from practice to practice. Consequently, T2 uptake in primary care was reported by the developers to be unremarkable. T2Developer revealed that there was lack of integration between T2 and the disparate systems used by GP practices.

Some GPs have the skills to do stable prostate cancer follow ups and some are less comfortable. As GPs are organised in clusters, it is challenging to agree on the logistical and organisational issues such as who is going to head up the service in that area, where will the data reside, and how to filter patients into that particular service. The successful ones we have had are where the organisation is acting as a hub and there is a structure in place where patients actually merge to one particular location, and they might have several patients from various outlying GPs. ... That problem had been solved already before we got there so that's good, but [it is] more complicated in some situations.

T2Developer

However, T2LeadConsultant argued that the main reason for T2's failure in primary care (in the region where it was developed) was that Band 5 nurses who were originally intended to use it did not have the skills to manage this patient group, even with T2 support. Consequently, the Band 5 nurses reportedly referred many patients back to hospital based consultant uro-oncologists. He also noted that there was widespread speculation about the proposed changes to the structure of Primary Care Trusts during



this period, which added uncertainty for decision makers regarding T2 uptake. As a result, the developers explored new opportunities to also implement T2 in hospitals. T2Developer noted that hospitals had many advantages over primary care such as high patient volumes to justify IT investment for decision makers, better economies of scale, centralised decision making and easily accessible patient data. The study Trust was the first to adopt T2 for use in a hospital setting.

In the hospital, you have a number of consultants that are on board and they have a few meetings and agree to send patients to clinic. Hospitals provide patient volume and good access to the data, unlike in community settings where data are spread around several servers, making it more challenging. And in some situations nurses may have to key information in again. ... There is equal interest from both GPs and hospitals but there are probably more economies of scale within the hospital environment and possibly financially, it may be difficult to put forward an argument in the community, and politically, it would appear to be more difficult on the face of it to justify the system in a community setting but compelling arguments are being developed for community use.

T2Developer

## **6.4 Implementation of T2 in the study Trust**

The initial plan was to implement T2 in the study Trust at the beginning of 2011. However, the implementation was delayed for several months because the study Trust's ICT department had concerns about information governance. This was mainly because patient data were going to be kept by T2 developers rather than on the study Trust's own servers. Additionally, the ICT department did not feel that T2 provided the value for

money that had been claimed by the developers. However, the decision had already been made to adopt T2 and funding had been agreed with the commissioners. T2 was eventually implemented at the end of 2011. However, the implementation was not supported by the study Trust's ICT department and the planned integration of T2 with the study Trust's existing clinical results reporting system did not occur. T2Developer revealed that that integrating T2 with existing legacy systems was not challenging itself but getting access to the study Trust's ICT infrastructure was. He noted that this was mainly due to information governance concerns by adopting organisations and lack of commitment to the project by key decision makers (especially the ICT department).

Making one computer talk to another is relatively straightforward. The problem is not technical... it is dealing with the hospitals in general, protocols, the red tape, getting to the data... I understand the many good reasons for data security... The biggest problem is securing the hospital IT resource, obtaining information governance clearance, to release information or make sure that information will be used in a safe manner... agreeing with IT departments a method that is mutually acceptable, especially across Trusts because we are having to get two discreet teams jelling together... it's quite a challenge.

T2Developer

Training to use T2 was provided to the uro-oncology nurses by the developers over two half-day sessions. T2Nurse2 described the training as "basic but adequate" and noted that T2 processes were largely similar to the study Trust's clinical pathways that were also based on NICE guidelines. Newly referred patients were to be seen through the T2 telephone clinics, while existing patients remained in the "old clinics". However, the lack of integration with existing legacy systems meant that the nurses had to log into T2

separately and manually enter demographic data and clinical results. T2 generated clinic letters also needed to be downloaded and then uploaded onto the study Trust's clinical results reporting system which initially caused a challenge because of incompatibility issues. The lack of integration also resulted in duplication of many of the clinic processes. Preparing for and managing the clinics was reported to be taking a lot longer than anticipated and consequently fewer patients were being seen in comparison with the previous system. This meant that the time savings which had been promised by the developers were not being realised. In some instances, the nurses had to override T2 recommendations. Examples include instances where T2 recommended that the nurse should order blood tests or scans or refer the patient to the uro-oncology consultant. In most of the cases, the patients were already known to have abnormal results, without necessarily implying that their condition was worsening.

It is more time consuming, obviously, because you are navigating three systems [within one clinic] and you have to make sure to find time before the clinic to input data from the other [legacy] systems into T2. But with time you get used to it, and the fact that you are not dictating letters for GPs makes it easier in the end, but while you are running the clinic and while you are speaking to the patient you are having to navigate three different systems and that takes time but eventually you get used to it, whether it a good thing or not, but then you don't see it as a hindrance in the long run because you have gotten used to the process.

T2Nurse1

T2Nurse1 also argued that background knowledge of uro-oncology nursing was essential because T2 was not capable of making such distinctions.

You couldn't rely on the system [T2] to generate all the answers for you or just accept everything it recommends... it will generate a lot of medical issues but sometimes you need to have the clinical knowledge to deal with the practical issues... clinical knowledge such as the treatments they have had and their side effects, reading the results and appropriate follow up...

T2Nurse1

## 6.5 T2 evaluations

A number of evaluations were performed during T2's development and following its adoption by the study Trust. These evaluations are summarised in Table 6.1 in chronological order. These evaluations included clinical efficacy and patient safety testing during T2's development, post market surveillance following implementation and patient satisfaction assessments.

<b>Evaluation</b>	<b>Method</b>	<b>Timing and duration</b>	<b>Evaluators</b>	<b>Evaluation purposes</b>	<b>Target audiences</b>	<b>Results</b>	<b>Uses of evaluations</b>
<b>Validation of software rules</b>	Clinical validation	During T2's development (2006-2009)	T2 developers and clinical researchers	To test system's alert levels and safety parameters	MHRA and Commissioners	Rules added to algorithms to improve patient outcomes	To obtain MHRA registration and show system's robustness to commissioners
<b>Clinical evaluation of the system</b>	Comparison of human and computer decisions	During T2's development (2006-2009)	Clinical researchers and system's developers	To evaluate system's accuracy in making guideline adherent management plans compared with humans	MHRA and developers	System found to accurately follow guidelines and recommend appropriate management plans	To obtain MHRA registration and show system's robustness to commissioners
<b>Pre T2 adoption evaluations</b>	Assessment of presentations and developer's literature	Prior to T2 adoption in study Trust (2011)	Lead Consultant at study Trust (T2LeadConsul)	To assess the merits of T2 before adoption	Commissioners and Trust Board	T2 found to be suitable for the Trust's requirements	T2 implemented in the study Trust in 2012
<b>Post market surveillance</b>	Audits	Performed over a 6 month period following implementation (from January 2012 to June 2012)	T2 developers	Inspecting system's accuracy and checking patient safety	MHRA and commissioners	System was been found to be accurate	Obtaining MHRA approval and showing the T2' robustness to commissioners
<b>Measuring patient satisfaction</b>	Survey	Performed 9 months after implementation over a 2 month period (from September to October 2012)	Clinical researchers	To establish patient satisfaction with telephone clinic and internet access	Researchers and commissioners	Majority of patients satisfied with telephone clinic. Few patients had access to the Internet	Clinical researchers planning to undertake further evaluation studies

Table 6.1 T2 evaluations in chronological order

### 6.5.1 Validation of software rules

From 2006 to 2009, T2Developer noted that they attempted to validate the T2 software rules through “iterative tests” to check T2’s alert levels, patient safety aspects and adherence with NICE guidelines.

The system was ready after 18-24 months and the first operational system installed after 3 years. The remainder of the time was used for testing rigorously with lots of iterative tests over various sets of data to make sure that the system was safe.

T2 Developer

The tests were performed by a team of software developers, clinicians and academics from a regional medical school. The evaluators used contrived clinical information to test T2’s responsiveness to different clinical scenarios that were typically encountered in practice. This was followed by similar tests on sets of retrospective anonymised clinical data that were sourced from existing electronic patient record systems at the developing Trust. In this case, real patient profiles, demographic data and co-morbidities were utilised, together with actual PSA levels, other blood results and radical treatments performed. T2Developer noted the high level of precision that was required to write clinical decision support software in comparison with other industries. He argued that this was especially so for intelligent systems like T2 that provides advice and guidance for managing patients.

Translating the paper algorithms into software is relatively challenging because there is zero margin for error... the biggest challenge was to make sure that we had robust testing... developing CDSS requires specialist clinical knowledge and understanding of the safety requirements, an intimate knowledge of how hospitals and community settings work, as well as guidelines among other essential issues... Testing regimen is

extraordinary and lengthy; clinical trials, peer reviews, consultant feedback and review to cover loops... strict adherence to NHS standards.

T2 Developer

All the T2 software rules were found to be adherent with NICE guidelines [NICE, 2008]. The targeted audiences of the validation tests were mainly the MHRA and NHS commissioners. These tests were considered a key factor in making the decision to adopt T2 by the study Trust. T2LeadConsultant noted that T2's validation was more rigorous than the other systems they had looked at, thus considered to be more robust and safe to use. The reports of these tests were not made available to the researcher because the developers considered them to be commercially sensitive.

### **6.5.2 Clinical evaluation of the system**

A team of academic researchers together with the T2 developers undertook a clinical evaluation of the system over a three-year period from 2006 to 2009. The objective of the evaluation was to test T2's ability to follow NICE guidelines and generate clinically accurate recommendations for managing patients with stable prostate cancer. The evaluators compared outcomes generated by T2 with those of three clinicians; two consultant urologists and a urology specialist registrar. 100 sets of retrospective anonymised patient clinical data were obtained from the existing electronic patient record system at the developing Trust. They included demographic data, relevant history, co-morbidities, blood results, symptom scores, diagnoses and radical treatments undertaken. The data sets were given to the clinicians to individually assess each case and make a decision regarding how they would manage the patient. The same data sets were also input into T2. Management plans generated from the assessments were in the form of codes. The codes were collated and

comparisons were made between the recommendations that were made by the clinicians and those that were generated by T2. A total of 236 codes were produced for 100 patients but there was complete agreement between the two sets of codes for only 46 of the 100 patients. An expert panel of three consultant urologists reviewed the discrepancies and found that they were mainly related to different interpretations of data by the participating clinicians. The expert panel concluded that in every one of the cases, T2 had adhered to the NICE guidelines and had recommended an appropriate management plan. T2DevelopingConsultant noted that the clinicians missed some “significant” clinical information in their assessments. He noted a 10% human factor error rate, which he argued was common in human investigations. Some of the errors are noted below:

- missed abnormal liver function tests while patient was on hormone therapy which could have potentially resulted in liver failure
- wrongly diagnosing prostate cancer recurrence
- multiple missed blood tests
- patients brought back to clinic too early
- patients brought back to clinical too late, e.g., patients showing signs of recurrence following surgery
- missed high creatinine and potassium levels which could have potentially resulted in renal failure

Also in some instances, the participating clinicians did not follow NICE guideline recommendations. For example, NICE guidelines recommend that patients should be offered a flexible sigmoidoscopy five years after radiotherapy. A reminder is embedded into the CDSS to ensure that the test is booked when the patient attends a routine clinic visit. The clinicians



were given an opportunity to review their own codes against recommendations made by T2. The revisions resulted in agreement on a further 44 patients. The expert panel reviewed the discrepancies for the 10 remaining patients again and confirmed that the T2 management plans had fully adhered to NICE guidelines. One of the T2 software rules was amended in the clinical pathway to reflect the expert team's findings to improve patients' management plans. After the rule change, the 100 cases were reprocessed by the system and resulting management plans were compared with suggestions that were recommended by the expert panel. Again, T2 adhered completely to the updated guidelines so the evaluators were able to conclude that T2 followed the defined algorithms and generated management plans that met the NICE guidelines (NICE, 2008) requirements. The targeted audiences for this evaluation were the MHRA, T2 developers and NHS Commissioners. T2Developer revealed that the T2 project team faced various challenges in undertaking the clinical evaluations, particularly getting uro-oncology experts to participate in pre-implementation evaluations and to validate the system.

We are fairly fortunate to have a Consultant Urologist as Medical Director [T2DevelopingConsultant]... otherwise getting expert knowledge, their time... borne out when we tried to get the expert panel together to look at the data. They are such busy people who can't even spare a few hours... fortunately, we secured 3 fellow Consultant Urologists... if you weren't in the privileged position in a hospital of having a Consultant involved in the project wholeheartedly it would never get off the ground, which makes it doubly difficult to do what we do...

T2Developer

### 6.5.3 Pre-T2 adoption evaluations

At the beginning of 2011, the T2 developers delivered a PowerPoint presentation to T2LeadConsultant, the divisional service manager and his fellow consultant urologists at the study Trust. The presentation covered aspects such as how T2 was developed and the evaluations that had already been carried out by the developers. The purpose of the presentation was to enable the developers to outline the potential benefits that would be achieved by the study Trust from implementing T2. For the study Trust, the purpose of the presentation was to assess whether T2 was suitable for their needs in comparison with two other CDSSs that T2LeadConsultant had already assessed. The first two systems that were considered were developed by consultant urologists in other NHS Trusts. They were both deemed too small for the study Trust's requirements. Additionally, they had not undergone satisfactory clinical validation to ensure patient safety. T2LeadConsultant selected T2 for a number of reasons. First, the Medical Director of the company that developed T2 (T2DevelopingConsultant) was known to him and they had previously collaborated on other projects. As such, he had confidence in T2 because he trusted his peers who had developed it, noting "we know them (developers), so we believed them. If it had been a purely commercial product, we would have been wary of what they were trying to sell us..." Additionally, the T2 Medical Director was also a senior consultant urologist at the developing Trust and he was using T2 for his patients. T2LeadConsultant noted that the urology department at the developing Trust was in many ways similar to the one at the study Trust. As such, he expected T2 to work in the study Trust in the same way as it had done in the developing Trust.

They [T2 developing Trust and their commercial partners] were looking for site to roll it [T2] out. We have previously collaborated with them [T2 developing Trust] and we knew the guys who developed it and it seemed to be working well for them. They had a similar unit to ours, so we thought we would give it a go, particularly as the implementation costs were going to be paid by their commercial partners.

T2LeadConsultant

Also, unlike other systems they had considered, T2 was a commercial product that was registered with the MHRA and had CE marking. This further proved to the study Trust that the necessary clinical and technical validations had been carried out in accordance with regulatory requirements.

We had looked at another system developed by a Consultant in the South West. It had not been commercialised or tested [and had] grown organically out of the Consultant's own special interest in-house and only had capacity for 500 patients. Another system, again in-house, was not ready for implementation elsewhere ... The other thing we liked about T2 was the CE mark; that and the fact that we trusted the guys who developed it.

T2LeadConsultant

T2LeadConsultant noted that if he were to carry out T2 evaluations himself, he would "do exactly the same as they had done [T2 developers]... and there was no point to do it [evaluations at the study trusty] because they [developers] had already done it", i.e., comparing the decisions of participating clinicians with those generated by T2. The study Trust was also reassured that there would be continued technical support following adoption. Furthermore, T2 allowed patients to be seen in nurse-led clinics, which the study Trust

deemed to be more cost-effective than consultant-led clinics. Finally, adoption and implementation for T2 was fully funded by the Primary Care Trust and one of the pharmaceutical companies that had co-developed it using the 'Practice Based Commissioning' model which was being championed by the T2 developers. There was no requirement for the study Trust's decision makers to submit business plans to commissioners to justify the required investment in its adoption. T2leadConsultant also noted that if they were to carry out any evaluations, they would perform similar evaluations to the one carried out by T2 developers.

The best evaluation is comparing Consultants' decisions and what the computer would have done but that takes too much time. The attraction for us is that they [T2 developing Trust] had already done that so we were happy to take their word and we were only happy to take their word because we knew them from before and we trusted them, whereas if it was a completely commercial outfit that would make us more suspicious as to what they were just trying to sell to us.

T2LeadConsultant

#### **6.5.4 T2 post market surveillance audit**

The T2 developers reported undertaking a "post market surveillance audit" for 6 months following its implementation in the study Trust (from January 2012 to June 2012). The purpose of this audit was to inspect the accuracy of T2 recommendations and to ensure patient safety. The developers were required by the MHRA to collect data and report on T2's safety following implementation. However, it was up to individual developers to implement their own methods of surveillance and define their own testing parameters. T2 developers inspected only 10% of the outcome letters generated from T2 clinics in the first 6 months at

the study Trust. The recommendations from the outcome letters were checked for compliance with NICE guidelines (NICE, 2008) and local clinical pathways. The developers also reported using a computerised methodology to test for T2's accuracy. They reported that there had not been any patient safety concerns since T2 was implemented in the study Trust. Interestingly, none of the interview participants at the study Trust (including T2LeadConsultant) were aware of this audit. Requests by the researcher for the post market surveillance audit report were unsuccessful because it was considered by the developers to be commercially sensitive.

### **6.5.5 T2 patient satisfaction survey**

From September 2012 to October 2012, clinical researchers from a medical school that was linked with the study Trust undertook an exploratory survey of the patients who had had consultation in the T2 outpatient telephone clinics at both the developing and study Trusts. The primary objective of the survey was to assess patient satisfaction levels with the T2 telephone clinics. They also sought to establish how many patients had access to the Internet, and whether these patients would be willing to use a personalised and confidential web based portal to give the T2 nurses information about their symptoms and concerns and to choose from a range of health topics that were important to them or wished to discuss with the nurse prior to the T2 telephone clinics. 25 questionnaires were returned from the study Trust and 29 from the developing Trust. However, it was unclear what the response rate was because the survey was not fully documented or publicised. Figure 6.3 shows a comparison between the respondents from the developing and study Trusts in relation to the T2 telephone clinics.

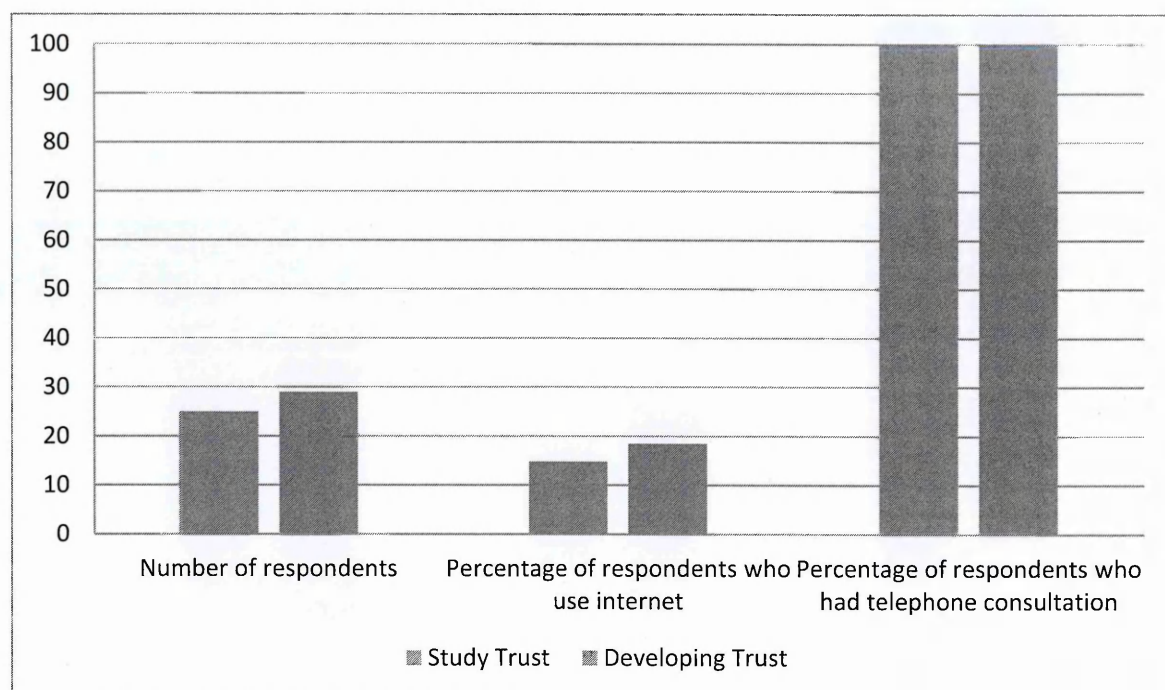


Figure 6.3 Respondents who participated in T2 telephone clinics (showing comparison between developing and adopting Trusts)

Overall, T2 telephone clinics were popular with the vast majority of respondents in both Trusts. 60% of the respondents in the study Trust had had at least one telephone consultation with the nurse. 75% of respondents from the developing Trust had had two or more telephone consultations with the nurse. The majority of patients in both Trusts (over 85%) reported their experience of the telephone consultation; ability to discuss their symptoms with the nurse; their understanding of the information that they were given; and how it helped them to manage their health as ranging from good to excellent. However, the majority of the patients in both Trusts, especially those aged 75 and above, had no access to the Internet at home or elsewhere. Only 20.8% of the respondents reported that they would find the internet portal to give the nurse information about their symptoms prior to the telephone clinic useful, while 15.1% reported that it would not be useful and another 15.1% did not know. 49% of the respondents noted that it would be inapplicable to them. The majority of respondents (62.3%), most of whom had been diagnosed within the last five years, did not

feel that submitting their clinical information online prior to the clinic would be useful. 60% of the patients who reported that the portal would not be of any use or would not be applicable did not have internet access. This group of respondents was mainly aged 75 and above. These findings were important because NICE guidelines (NICE, 2014) strongly recommended both telephone and secure electronic follow up (i.e., via the Internet) instead of the traditional face to face outpatients' clinics in hospitals. The evaluators concluded that a more thorough evaluation was warranted. In 2014, plans were underway to prepare a proposal for a collaborative project between the university and both Trusts. T2LeadConsultant indicated that the evaluators were most likely to undertake an audit rather than "actual research". It was unclear why the evaluators were choosing an audit over other forms of evaluation. The results of the exploratory study are shown in Figure 6.4. This figure shows the number of respondents at the developing and study Trusts who thought that using an electronic portal to report their symptoms before attending T2 telephone clinics would be useful.

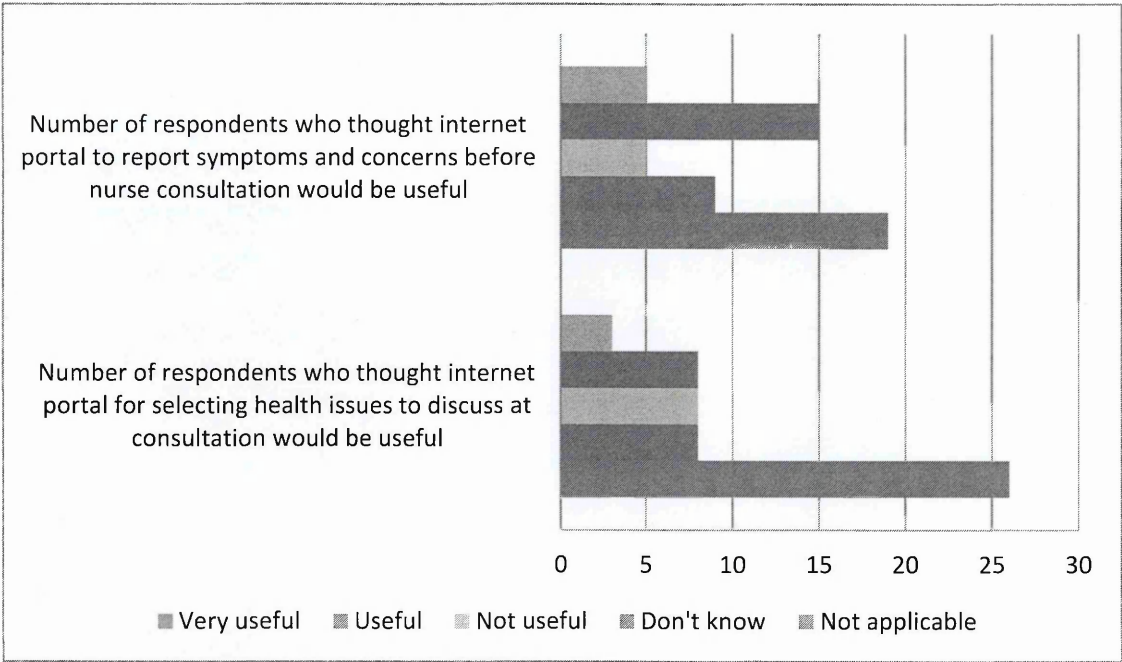


Figure 6.4 T2 exploratory survey results showing comparison between respondents from the developing and study Trusts

## 6.6 Discussion

In this section, the evaluations that were carried out in relation to T2 are examined in the context of the key aspects of evaluation that were identified in the literature review. This is done to establish the extent to which these evaluations were similar or different from those undertaken in other healthcare settings, as well as establishing missed opportunities.

### 6.6.1 Purposes of T2 evaluations

The majority of T2 evaluations were performed by the developers during its development in collaboration with the project partners. The purposes of these evaluations were primarily to check T2’s technical and clinical accuracy, test patient safety parameters, guideline adherence and regulatory requirements. Although the developers made wide ranging claims about the economic benefits, improvement in performance and patient outcomes, none of the reported benefits were specifically addressed by the evaluations performed. Improvements in user



performance and clinical efficacy were assumed rather than proven by evaluations. T2LeadConsultant acknowledged the need to establish these benefits, noting that “it would be great if I could tell you how many slots we freed up... how many more patients we saw...” but revealed that they that did not have the financial and human resources to perform such evaluations.

There is no evaluation at all [at the study Trust]. They [the T2 developers] had pretty much the only system that was out there so our choice was to use their system or develop our own. I haven't got the expertise to do it and someone was going to pay for this one. ... There are a lot of patients waiting for appointments, nobody knows how many out there. If things are going well, the secretaries will just book them and if it's unwell, there is a backlog... At no point am I aware of exactly how big my outpatient population is or how long they are waiting over what I have suggested. In a perfect world I would be aware of that data and I would be able to say how many slots I have freed up every week. I do know they are seeing patients every week because if they weren't, I would be having to see them. What the benefit is, I have no idea...

T2LeadConsultant

The patient satisfaction survey was the only evaluation which looked specifically at patient outcomes. However, it was a small, partially documented exploratory study which had little impact on decision making in the study Trust. It did nonetheless dispel some of the assumptions which were held by both the developers and decision makers at the study Trust. For example, in line with NICE guidance, the developers highlighted that this patient group preferred to be followed up in primary care settings and over the telephone or internet rather than in face to face outpatients' clinics. However, the survey results showed that while

patients gave positive feedback about the telephone clinic, they were less keen on using the internet or related tools to communicate with the T2 nurses regarding their symptoms. T2Nurse2 also reported that most patients preferred face to face clinics compared with the telephone clinics because it gave them a chance to maintain human contact and also to discuss other problems that they may have. The post market surveillance audit was primarily focused on satisfying MHRA requirements and CE marking. However, the developers were keen to label these audits largely as evaluations of patient safety. Although effectively these longitudinal data contributed towards patient safety in as far as guideline adherence was concerned, the primary motivation for the developers was the commercial value of having an MHRA approved and CE marked product that they could sell to other NHS Trusts.

### **6.6.2 Approaches and methods used to evaluate T2**

The two evaluations that were carried out by the developers used quantitative clinical validation methods that are commonly used in the NHS. These methods were primarily used to test the systems' robustness, usability and adherence to clinical pathways or guidelines. The key variables that were measured include system accuracy, system and patient outcomes and user performance. It could be argued that T2 developers carried out these evaluations because they knew that their peers who are responsible for commissioning T2 in NHS Trusts would approve their methods. Indeed, T2LeadConsultant acknowledged that the validations that were carried out by the developers were an important factor in his adoption decision and that he would have used similar methods to evaluate T2. However, clinical validations and audits do not explain the effects of T2 on the environment where it was introduced. Examples of these effects include changes to targeted users' roles, shifts in responsibilities and existing workflow. Indeed the system might prove accurate and efficient in a controlled

environment using contrived data but may fail to achieve the same results in a real clinical setting where various competing priorities are at play. For example, T2Nurse1 reported that additional jobs that were undertaken simultaneously with T2 clinics could not be coded, such as unexpectedly long clinic visits, dealing with related clinical problems which patients presented with and circumstantial clinical examinations or referrals that were not always “picked up” by T2. T2Nurse1 also reported having to override some T2 recommendations, noting that it was “too sensitive to the guidelines” and did not consider the nurses’ expertise and experience.

I am not particularly aware of their thought processes [at the time of adoption], but looking at it from where we are, it would be helpful for a person coming in on a developmental role [not for an experienced nurse]... They would still need training to run the clinic but it would be interesting to see whether they would be able to navigate the systems without as much training or [prostate cancer background] knowledge.

T2Nurse1

However, T2Nurse2 argued that regardless of their expertise and experience, T2 was an important “safety net” which ensured that they followed the NICE guidelines fully and did not miss any important symptoms or tests regardless of whether it was used in primary or secondary care settings.

Introduction of the system [T2] was an improvement in identifying and minimising risks in terms of decision making. With risk management, it is very important that you are able to identify if there is any element of risk in the way you manage the patient. The system [T2] prompts you. It’s not something that you wouldn’t think about but it makes you look at other avenues, because it looks at certain aspects such as bone

pain, frequency, erectile dysfunction, whereas you may fail to ask the specific questions, and it makes sure you are asking the right/relevant questions which will lead you to make an informed decision. You can never have too many reminders for a condition like prostate cancer - as long as it reminds you about something that will benefit patients, it's a win-win in terms of minimising patient risk and improving the quality of care. ... The system itself promotes a high standard of care, reduces variability; we are asking the same questions, the mandatory fields, that's a benefit in itself. I find it very beneficial but obviously, with any system, people will have different opinions.

T2Nurse2

Both T2DevelopingConsultant and T2LeadConsultant also argued that there was no guarantee that the specialist nurses were following the guidelines appropriately or that their clinical decision making was always accurate.

Guidelines cannot tell you everything. The attraction [of T2] is that changes in practice can be implemented by making necessary changes to the algorithm. ... It makes it easier to change things as patients are managed centrally. ... The algorithm has to be capable of coping with any eventuality. There are times when a senior nurse can make the right decision but junior nurses may not be able to do so, so the computer thus makes perfectly reasonable decision. ... To my mind, it [T2 nurses overriding the system] doesn't happen often to warrant changing the algorithm.

T2LeadConsultant

However, the evaluations undertaken by the developing Trust looked at a system designed for use by Band 5 nurses in primary care, rather than senior specialist nurses (i.e., T2Nurse1 and T2Nurse2) who worked autonomously in hospital settings. The specialist nurses may not

require the same level of decision support that would be appropriate for junior nurses (e.g., Band 5 nurses) although clinical decisions should be the same for both groups of nurses. However, the study Trust may not have achieved the purported financial benefits because these nurses would cost more than junior nurses. The developers explained that senior nurses were used for patient safety reasons while they were waiting to being fully registered with the MHRA, i.e., submission of post market surveillance audit data. This may suggest that T2 was not suitable for primary care settings by junior nurses as originally intended. However, its usefulness for nurses who had a higher level of proficiency about urological cancers was questionable and in this case had not been evaluated. Although both T2Nurse1 and T2Nurse2 valued T2's structured format and clinic outcome letters, they disagreed regarding its added value to their work. T2Nurse1 viewed T2 as an "unnecessary" system which instead of helping, created more work for them while T2Nurse2 argued that regardless of the "teething problems" which had surrounded its introduction, T2 was an invaluable CDSS that would ensure patient safety and improved clinical decision making.

Probably the full potential of the system will be noticeable once fully integrated. ...

Whenever you introduce something new in a big organisation where there are other things in place, you have to learn about the new system, what aspects you want to marry. Obviously there are two or more systems so you want to reduce duplication in the system... [Asking] what can be added from the new system to support the old

system can take a while. ... I would not say delay [integration], I would say getting a better understanding before establishing what can be done.

T2Nurse1

Additionally, T2Nurse1 noted that they had to summarise clinic visits in patients' medical case notes because they were unsure whether the printed T2 outcome letter alone would suffice as a medical record and they wanted to "cover their backs". These concerns exposed grey areas around professional duty of care, ethics and legal imperatives of novel technologies in tightly regulated healthcare environments.

### **6.6.3 Contextual and organisational issues that affected T2 evaluations**

None of the evaluations performed looked at the contextual and organisational issues such as the differences between the originally intended T2 implementation context and the study Trust. There was an assumption by project leaders that T2 would also work in the study Trust because of its similarities with the developing Trust. T2Developer revealed that having T2DevelopingConsultant as their Medical Director and also being a senior consultant at the developing Trust was invaluable to the team. This dual position carried significant influence on T2's adoption and evaluation in the developing Trust. Indeed, the clinicians at the study Trust did not have the same influence as those in the developing Trust as evidenced by the lack of support from key decision makers such as the ICT department. Also the assumption by T2LeadConsultant that the urology departments in both Trusts were similar only took into consideration technical and structural aspects of T2 and the NHS in general. Indeed, this assumption effectively removed the evaluation agenda from the adoption process without involving other stakeholders in the study Trust. It would appear that the complex interplay between the actors involved, political, social and other organisational factors were ignored.

The literature review (see chapter 2) suggested that CDSS evaluations needed to consider the interdependencies between these factors and also coordinate all the key stakeholders in any evaluation activity. It appears that the study Trust's ICT department were not involved in the T2 adoption process until the time of implementation. Failure to involve the ICT department earlier may have been a missed opportunity to assess T2's suitability for the study Trust and its compatibility and integration with existing systems before the adoption decision was made. There may also have been an opportunity to set the boundaries for evaluation activities in line with the Trust's strategic goals. The uro-oncology nurses were also not involved in the T2 adoption decision. However, T2 brought about significant changes in their ways of working. For example, they were now spending more time inputting data, and consequently seeing fewer patients per clinic in comparison with the previous system. These issues were not addressed in any of the evaluations. T2Nurse1 noted that they did not raise these and other operational issues because the adoption decision had already been made and they "just got on with it".

Looking at the management point of view, they have a different idea of what the system is supposed to do. ... When you are actually working with it, you have a different point of view. ... So from a management point of view, I understand that somebody less qualified can run the clinic with support from the system rather than getting to another person, so that's fine.

T2Nurse1

#### **6.6.4 Barriers to T2 evaluations**

There were many barriers to T2 evaluation. The main barrier was that the study Trust did not see the need to undertake any evaluations at all. The primary reason given by

T2LeadConsultant was that the evaluations that had already been carried out by T2 developers were satisfactory. This was mainly because he had previously worked with T2DevelopingConsultant and was prepared to “take his word” regarding the robustness of these evaluations. Additionally, he noted that T2’s registration with the MHRA and CE marking had “reassured” him of its robustness. It would appear that the decisions both to adopt and not evaluate T2 following adoption were based on trust of the developers rather than the robustness of evaluations undertaken as reported. However, as noted above, these prior evaluations looked at the use of T2 in a completely different context and with different users. It is therefore difficult to generalise the benefits or appropriateness of these evaluations to the study Trust’s context. T2Nurse2 argued that the study Trust needed to give T2 time to “settle” before undertaking any evaluations to assess its effectiveness. However, by that time T2 had already been in use in the Trust for over a year and had become a central part of the management of stable prostate cancer patients. Other barriers to T2 evaluations noted included the lack of time and resources, the lack of motivation to evaluate due to organisational changes and the top down nature of the departmental structure. These barriers were closely linked with the decision made by T2LeadConsultant not to evaluate at all in the study Trust. At the time of T2 adoption there were many structural changes to NHS funded primary care services. The local Primary Care Trust reorganised into two Clinical Commissioning Groups, which subsequently evolved into three within a year. T2LeadConsultant argued that it became unclear who was responsible for funding and monitoring the performance of T2 because many of the key stakeholders who were originally involved in its commissioning were now focusing on different priorities or had moved on completely.



I have heard that most people think they [the T2 developers] charge too much per patient, but we have negotiated a good tariff so that's not an issue for us. We may have a problem in future if the PCT (Primary Care Trust) look at the tariff and say "we can't pay this". ... The CCGs (clinical commissioning groups) have reorganised themselves in a different fashion. I would be amazed if anyone in the new structure would remember that it was negotiated; our manager who negotiated this is no longer with us. ... I don't know who he was talking to in the CCG. I know that's not the way it should be but how on earth do you determine whether it's working or not?

T2LeadConsultant

### **6.6.5 Benefits of T2 evaluations**

The main benefits of T2 evaluations were the satisfaction of regulatory requirements and facilitating its registration with the MHRA and CE marking. In turn, this gave T2 credibility as a robust CDSS that was safe to use and satisfied the organisational, ethical and professional duty of care for the study Trust. However, the limitations of these evaluations were not addressed. During T2's clinical evaluation at the developing Trust, it was noted that the clinicians missed some relevant clinical information in their assessments and in some instances did not follow NICE guideline recommendations. It could be argued that regardless of how experienced or senior a clinician is, there is no guarantee that they will follow the guidelines fully or deliver safe and appropriate patient care. The same could also be said for the hospital-based specialist nurses who felt that T2 was developed for junior nurses who required that level of guidance and clinical decision support. It could then be argued that T2 promoted best practice, regardless of the users' seniority or lack of experience. However, there was a gap in the evaluations that were undertaken because they did not look at how T2 would be used by hospital based specialist nurses and how such a CDSS would affect their

established working patterns. Also the originally intended use of T2 had significantly changed and it was unlikely that the originally “sold” benefits would be achieved in hospital settings. Clearly, there were many questions that had not been answered by the evaluations that were carried out and there was insufficient economic or cost-effectiveness evidence to support the continued use of T2 in hospitals as claimed by the developers.

### **6.6.6 Overview of T2 evaluations and missed opportunities**

The T2 evaluations did not address many important issues that were identified in the literature review. The evaluations performed by the developers were only concerned with the technical and safety aspects as required by the regulators. These evaluations were also focused on T2 use in primary care settings by junior nurses rather than by senior nurses in hospital settings such as the study Trust. Also the adopting Trust did not commission any evaluations at all. It was difficult to judge whether they achieved the expected benefits or if there were any benefits at all from adopting T2. By not formally evaluating T2, the study Trust missed an opportunity to establish the effects of T2 on the users and patients, clinical workflow and the wider organisation. There was a lack of motivation to evaluate and challenge assumptions about the effectiveness of T2. Also, the developers only performed formative studies that were necessary during CDSS development. A summative evaluation would have identified whether T2 had solved the problems that it was designed to solve. However, the study Trust accepted the benefits of T2 based on the evaluations reported by the developers. This was based mainly on trust of their peers and its links with NICE, MHRA and CE marking rather than robustness of evaluations. The gaps in evaluation mainly related to contextual and organisational issues and the range of evaluation approaches employed.

Key evaluation factors based on the CDSS evaluation framework	T2 evaluations carried out	Missed opportunities
Purposes of evaluation should cover a range of issues, rather than focus on single issues	<ul style="list-style-type: none"> <li>• Most evaluations carried out by developers</li> <li>• Primary purposes of evaluations were to test T2's clinical and technical effectiveness, guideline adherence and immediate patient safety aspects</li> </ul>	<ul style="list-style-type: none"> <li>• T2 fit for study Trust</li> <li>• T2's economic effectiveness</li> <li>• Patient benefits in a typical hospital setting</li> </ul>
Various approaches and methods of evaluation should be applied throughout the CDSS's lifecycle	<ul style="list-style-type: none"> <li>• Quantitative clinical validation methods</li> <li>• Assessment of guideline adherence using contrived and actual patient data within controlled environments</li> </ul>	<ul style="list-style-type: none"> <li>• 'Sold' benefits not tested in study Trust</li> <li>• No evaluations carried out in a hospital setting with experienced uro-oncology specialist nurses</li> <li>• Lack of integration with legacy system and T2 impact on workflow settings not evaluated</li> </ul>
Contextual and organisational issues should be considered in any evaluation activity	<ul style="list-style-type: none"> <li>• Study Trust contextual and organizational issues not evaluated formally</li> </ul>	<ul style="list-style-type: none"> <li>• T2 could have been evaluated in the context of the study Trust, rather than informal assumptions based on 'sold' benefits</li> </ul>
Benefits of evaluation should be made clear to all stakeholders	<ul style="list-style-type: none"> <li>• No formal evaluation of T2 benefits to study Trust</li> </ul>	<ul style="list-style-type: none"> <li>• 'Sold' benefits of T2 could have been evaluated at different stages of its adoption in the study Trust</li> <li>• T2 project team could have looked at workflow effects and benefits to patients</li> </ul>
Barriers to evaluation should be mitigated by widening the purposes, approaches and methods, taking into consideration contextual and organisational issues as well as highlighting the benefits of evaluation	<ul style="list-style-type: none"> <li>• No formal evaluation of T2 barriers</li> </ul>	<ul style="list-style-type: none"> <li>• Assumptions by decision makers that evaluations carried out by T2 developers would apply to study Trust hampered opportunities to assess the fit of T2 to study Trust</li> </ul>

Table 6.2 Opportunities that were missed by T2 evaluations

## Chapter 7 T3 track and trigger tool

This chapter looks at the evaluation of T3. It is an expert electronic track and trigger clinical decision support system (CDSS) that was used to monitor, identify and respond to deteriorating patients in acute hospital settings. The introduction section provides a broad overview of the key aspects relating to acutely ill medical patients and how T3 was intended to resolve these problems. Sections 7.2 and 7.3 look at the key national initiatives that were developed to manage this patient group in the UK. The study Trust context is introduced in Section 7.4, noting how T3 was implemented. The evaluations that were carried out for T3 are discussed in Section 7.5. The key aspects of evaluations undertaken are discussed in Section 7.6.

The participants who were interviewed for this case study are identified as follows:

1. T3 Lead Consultant – T3LeadConsultant
2. T3 Lead Nurse – T3LeadNurse
3. T3 Implementation Nurse – T3ImplementationNurse
4. Critical Care Outreach Sister – OutreachSister1
5. Critical Care Outreach Nurse – OutreachSister2
6. Ward Sister in Combined Specialties – WardSister1

## 7.1 Introduction

The National Patient Safety Agency (NPSA) reported that over 13 million patients were admitted to acute hospitals in England and Wales annually (National patient Safety Agency, 2007). They noted that although the majority of patients received safe and effective care, many were at risk of deteriorating and becoming critically ill during their admission, which in some cases resulted in avoidable serious adverse events or even deaths. Previous studies had shown that in most cases, patient deterioration was well documented but inadequate interventions were commenced and sometimes delayed (Sax and Charlson, 1987; Smith and Wood, 1998; Goldhill et al., 1999). Goldhill et al. (1999) found that patients who were transferred from the wards to intensive care units had higher levels of deaths compared with those who were admitted from operating theatres or emergency departments. Other studies found that although most of these deaths were predictable and preventable, physiological deterioration was not identified early on the wards and many patients received suboptimal care (Massey et al., 2009; McQuillan et al., 1998). Goldhill et al. (1999) highlighted the importance of early identification of patients at risk and ensuring that the appropriate interventions were commenced before they deteriorated further. Another study by Goldhill et al. (2004) concluded that the longer patients were on hospital wards before an intensive care unit admission, the worse their chances of survival were. They also found that nearly 25% of deaths following admission to intensive care unit occurred after discharge to the wards and many of these deaths were in patients who were previously considered to be relatively low risk (Goldhill et al., 2004).

In 2005, the NPSA received 484, 441 reports of patient safety incidents and 1804 deaths in NHS hospitals (NPSA, 2007). Most of these incidents and deaths were related to failures by healthcare workers to recognise physiological deterioration and commence appropriate interventions on time. Previous studies had shown that such failures often resulted in further physiological deterioration, which in turn led to admissions and prolonged stay in intensive care units, avoidable deaths and significant costs to healthcare organisations (Schein et al., 1990; Franklin and Mathew, 1994; McQuillan et al., 1998). The NPSA reported that NHS hospitals were now increasingly treating older patients who often presented with more complex and acute problems with additional co-morbidities (NPSA, 2007). Furthermore, resource limitations meant that only a few patients could be adequately monitored and treated in intensive care units (NICE, 2007).

## **7.2 Interventions for acutely ill medical patients**

This section will discuss the critical events, interventions and research studies that influenced the development and adoption of T3. The Early Warning Score (EWS), also known as track and trigger tool was a bedside physiological evaluation tool (Subbe et al., 2001). It was a paper-based tool that was developed to secure the timely presence of appropriately skilled clinical help by the bedside of those patients exhibiting physiological signs compatible with established or impending critical illness in acute clinical environments. Patients' blood pressure, temperature, respiratory rate and heart rate were recorded and a risk score was calculated based on whether they were within or outside acceptable ranges. The higher the score, the higher the likelihood of the patient deteriorating physiologically and potentially requiring admission into the intensive care

unit. Additional clinical parameters such as urine output and patients' alertness levels were later added to create a Modified Early Warning Score (MEWS) based on local protocols or clinical specialties. Gardner-Thorpe et al. (2006) demonstrated the value of the MEWS in appropriately identifying patients who would benefit from being admitted onto intensive care unit beds. Similarly, in a prospective cohort study, Subbe et al. (2001) validated the MEWS to identify medical patients at risk and its feasibility as a screening tool to trigger early interventions or determining patients' suitability for intensive care unit admission. They concluded that the MEWS was a simple and effective bedside tool that could be effectively used by nurses in busy clinical areas. NICE guideline 50 for acutely ill medical patients (NICE, 2007) recommended the use of electronic track and trigger systems based on the MEWS across the NHS. However, the NPSA reported many shortcomings in NHS systems which made it difficult to monitor, prevent, detect and treat physiological deterioration even though the MEWS had been widely adopted in the majority of NHS hospitals (NPSA, 2007). Hammond et al. (2013) also argued that despite widespread use of MEWS, clinical monitoring on the wards was often poor. They noted many inaccuracies in the recording of physiological vital signs and MEWS calculations, which in turn affected the detection of those patients at risk of deterioration. This in turn delayed interventions and resulted in needless overuse of limited resources. Since the early 2000s, many NHS Trusts invested in the development of in-house and acquisition of commercially developed electronic track and trigger systems that were based on the MEWS score. Some small scale studies concluded that electronic track and trigger systems were more accurate and efficient in comparison with MEWS based paper charts (Smith et al., 2006). Some studies (Gao et al., 2007; Smith et al., 2008a; Elliott et al., 2015) concluded

that electronic track and trigger systems reduced hospital deaths. However, they argued that many of these systems had not been formally evaluated. As such, there was limited information about their cost-effectiveness, effect on patient outcomes, resource allocation and use, and whether NHS organisations were benefiting from investing in these systems.

## **7.3 Management of acutely ill patients in the NHS**

Most clinical studies that were carried out from the late 1980's to the late 1990s found overwhelming evidence that acutely ill patients in hospitals were receiving suboptimal care. Since the mid-1990s, critical care services focused on improvement of patient outcomes and appropriate use of resources. Rowan et al. published a large study looking at the outcomes of patients in 26 intensive care units across the UK and Ireland (Rowan et al., 1993). They concluded that comparisons of death rates were misleading indicators of hospital performance. They also noted the heterogeneity of data collection methods and the need for case mix adjustments when comparing outcomes from different hospitals. Subsequently, the UK Intensive Care Society submitted proposals to establish a nationally coordinated comparative audit and research centre for critical care services. The Department of Health acknowledged the importance of the APACHE II study (Rowan et al., 1993) and emphasised the need for updated information on intensive care services. Financial support was provided for the establishment in 1994 of the Intensive Care National Audit and Research Centre (ICNARC). It was a separate organisation from the Intensive Care Society. The main aim of ICNARC was to provide an independent national resource to monitor and evaluate NHS intensive care services. Additionally, ICNARC



provided data to intensive care practitioners and commissioners on the most effective interventions and organisation of services. The immediate objectives of ICNARC focused on existing service provisions and practices in intensive care units across the UK. This led to the development of models to evaluate and monitor performance of intensive care units and evaluation of various aspects of service provision and patient outcomes. To that end, ICNARC required all NHS Trusts in England, Wales and Northern Ireland to submit monthly data on patient outcomes to its national Case Mix Programme. The Case Mix Programme was an audit of all adult general critical care units in England, Wales and Northern Ireland. Since ICNARC's establishment in 1994, over 1.5 million patients have been added to its database. These data are analysed and disseminated to critical care practitioners and commissioners to provide them with accurate outcome measurements and state of service provision. ICNARC established the National Cardiac Arrest Audit in 2008 in partnership with the Resuscitation Council (UK). The audit monitors and reports the incidence of cardiac arrests in hospitals and subsequent patient outcomes. ICNARC also established a Clinical Trials Unit alongside the national audit activities. The Clinical Trials Unit primarily carried out randomised controlled trials and cohort studies aimed at utilising collected data to improve critical care practice and guide policy development.

In 1999, the Audit Commission published its *Critical to Success* report (Audit Commission, 1999), which introduced the concept of Intensive Care Outreach Services. It recommended that NHS Trusts should agree on "danger signs" that would enable the identification of patients at risk of deterioration in acute hospital settings. The concept of Intensive Care Outreach Service was further developed in the Department of Health's *Comprehensive Critical Care* report (Department of Health, 2000). It recommended that

all NHS Trusts should set up Outreach Services staffed by senior critical care nurses and doctors. Their role was to support ward areas to identify and appropriately respond to deteriorating patients (McArthur-Rouse, 2001). They primarily used paper-based clinical observation charts accompanied by local escalation protocols based on the MEWS. However, various studies found little evidence of improvement to patient outcomes resulting from the introduction of MEWS (Gardner-Thorpe et al., 2006; Subbe et al., 2003). Other studies also concluded that the MEWS scores were being inaccurately calculated, which undermined the effectiveness of Outreach Services and resulted in suboptimal care being provided to patients (Smith et al., 2008a; Smith et al., 2008b; Prytherch et al., 2010).

In 2007, NICE published its clinical guideline 50, which focused on the recognition of and response to acute illness in adults in hospital environments (NICE, 2007). It recommended that healthcare workers should correctly measure and record physiological observations in line with existing MEWS-based clinical track and trigger systems. Emphasis was placed on full clinical assessment and tailoring of physiological observations and clinical management plans to individual patients. NHS Trusts were tasked with ensuring that healthcare professionals were appropriately trained to adequately assess and escalate care. Additionally, NHS Trusts were tasked with developing and implementing local strategies for tracking patients and ensuring appropriate responses to clinical deterioration. NICE guideline 50 also recommended that NHS Trusts should adopt electronic track and trigger systems to improve the quality of care provided. Some NHS Trusts had already introduced electronic track and trigger systems to manage acutely ill patients based on MEWS protocols in the early 2000s. However, they were mostly small standalone systems that were used in the units where they had been developed [NICE,

2007]. It was also up to individual NHS Trusts whether to use paper-based or computerised track and trigger systems to adhere with NICE guidance. In support of the recommendations, a study by Mohammed et al. (2009) concluded that handheld computerised systems improved the accuracy and efficiency of MEWS based electronic systems in acute hospital settings in comparison with traditional pen and paper methods. They also noted that the computerised systems were found to be acceptable to nurses.

Also in 2007, the NPSA published its fifth Patient Safety Observatory report (NPSA, 2007). The report highlighted the importance of the National Reporting and Learning Systems used in England and Wales to identify where patients were being harmed, rather than helped by healthcare interventions. It also identified that too little attention was being paid to data collected from these safety reporting systems in order to learn from these experiences to improve patient safety. The report noted the existence of systemic failures in NHS hospitals, which resulted in healthcare professionals failing to identify and mitigate acutely ill patients. These failures resulted in further deterioration, clinical complications, severe disability, prolonged hospitalisation and sometimes avoidable deaths in acute hospital settings. The NPSA complemented the NICE guidelines and wider national programme to improve the outcomes of patients in acute hospital settings. However, the NPSA noted that because the National Reporting and Learning Systems was voluntary, it may have been subject to bias and data incompleteness. To mitigate these limitations, the NPSA also looked at information from litigation organisations and new research to gain a better understanding of the causes of adverse events and deaths in acute hospital settings. Furthermore, there was also evidence of poor documentation and communication between healthcare teams. This often resulted in delays in seeking

appropriately skilled help and interventions. Additionally, the NPSA also recommended that NHS Trusts should improve resuscitation procedures, communication amongst the MDT, appropriate training and development and ensuring that frontline staff had the appropriate equipment to perform their roles (NPSA, 2007)..

The Royal Medical Colleges and related collegiate and professional organisations also supported NICE guideline 50. They helped to set up ongoing training and development programmes to ensure that all clinical staff were competent in measuring and recording vital signs in line with NICE recommendations. In 2012, the Royal College of Physicians published the National Early Warning Score (NEWS) (Royal College of Physicians, 2012). It was based on MEWS and aimed to standardise the assessment of acutely ill patients across all NHS hospitals and to replace all existing early warning scores. The Royal Colleges of Physicians and Nursing, the National Outreach Forum and the NHS Training for Innovation developed online programmes to support NHS staff to use the NEWS appropriately. Many NHS Trusts, including the study site adapted the NEWS into routine clinical practice. However, in 2014, there were no published formal evaluations of the NEWS despite its widespread implementation across the NHS. Recent studies looking at the impact of early warning systems on patient outcomes found that most of the available evidence was of poor quality (McNeill and Bryden, 2013; Kolic et al., 2015). They noted that most studies reviewed used uncontrolled before and after designs which made it difficult to attribute causation to the early warning scores. They also noted that most early warning scoring systems were implemented with little control of confounding factors that may potentially affect outcomes.

## **7.4 Management of acute medical patients in the study**

### **Trust**

In the study Trust, acute medical patients were primarily admitted through the Accident and Emergency department. They were assessed using the triage system and then transferred to appropriate wards or theatres depending on their clinical needs, investigations that were required or confirmed diagnosis. A key part of the clinical assessment was the taking and recording of clinical observations such as blood pressure, temperature, respiratory rate and oxygen saturations. To ensure uniformity across the study Trust, a paper based Early Warning Score clinical observation chart was adopted in 2001. However, audits that were carried out at the time showed variable usage of these charts across the Trust. Manual calculations of MEWS scores were found to be inaccurate. The audits also found that the higher the MEWS scores were, the greater the failure occurrence and delays in instigating remedial action. Also in the early 2000s, the study Trust set up a nurse led Critical Care Outreach service. The key outcomes of this service were to identify and respond to acutely ill or potentially deteriorating patients in the hospital, prevention of admission/readmission to critical care, advanced clinical skills training for nurses, and induction and support for junior doctors. The outreach service ran from 8am to 6pm and linked ward areas, the hospital at night team, ward based medical teams and the critical care department. OutreachSister1 reported that prior to T3 implementation, it was difficult to coordinate the activities of the outreach team, keep track of referrals for deteriorating patients and ensure that ward areas requested help on time. There were also challenges in tracking patients between wards, to and from theatres

and ancillary services such as radiology. The Outreach Sisters also noted that it was difficult for medical teams to prioritise their work and track patients spread across the hospital.

I was involved in the introduction of paper-based MEWS charts in the Trust... it was difficult to decipher, hit and miss, illegible writing, ill patients often missed, hence T3 was a natural progression that allows instant access to scoring patients... you can see remotely what's going on instantaneously...improves communication with patients' team and critical care doctors...

OutreachSister1

Since the early 2000s, T3LeadConsultant had a dual role as the study Trust's Lead Consultant for acute services and developing its clinical technology strategy. He had previously worked on a national project with the Lead Clinician who co-developed T3. In 2006, the company that developed T3 was looking for a large Trust to partner with and implement T3 hospital-wide. At the same time, the Trust Board (study Trust) was keen to adopt a tracking system to manage patient flow in line with NICE guideline 50. T3LeadConsultant noted that he was more concerned with patient safety aspects of any new technology adopted. The study Trust saw an opportunity to develop processes around patient tracking using T3. Due to provisions in financing of the hospital facilities, a business case was not required at the time and the Trust Board approved the required funding for T3 adoption at the study Trust.

### **7.4.1 The technology**

T3 was developed by a third party supplier in collaboration with another NHS Trust (the developing Trust) from 2005 to 2007. Its software algorithms were initially based on the developing Trust's MEWS protocol. During this period, there were various concerns that acutely ill patients in NHS hospitals were receiving suboptimal care. Various initiatives resulted in the NPSA commissioned reports looking at clinical incidents, adverse events and deaths in hospitals (NPSA, 2007) and the NICE guideline 50 for acutely ill patients in hospital (NICE, 2007). Recommendations from these two documents were incorporated into T3 software algorithms to ensure that patients were appropriately monitored, assessed and the necessary interventions commenced before they deteriorated further (see Figure 7.1). These algorithms could also be adapted to local clinical pathways or changes to clinical guidelines as necessary. T3 was registered with the MHRA as a software medical device and achieved CE marking following the necessary clinical validation procedures at the developing Trust. The developing Trust reported that T3 implementation led to significant financial, operational and clinical benefits. However, the developers were looking for a large NHS Trust to partner with and implement T3 across the hospital. The study Trust was the first to implement T3 across the hospital in 2007. Figure 7.1 shows the problems that were expected to be resolved by implementing and using T3.

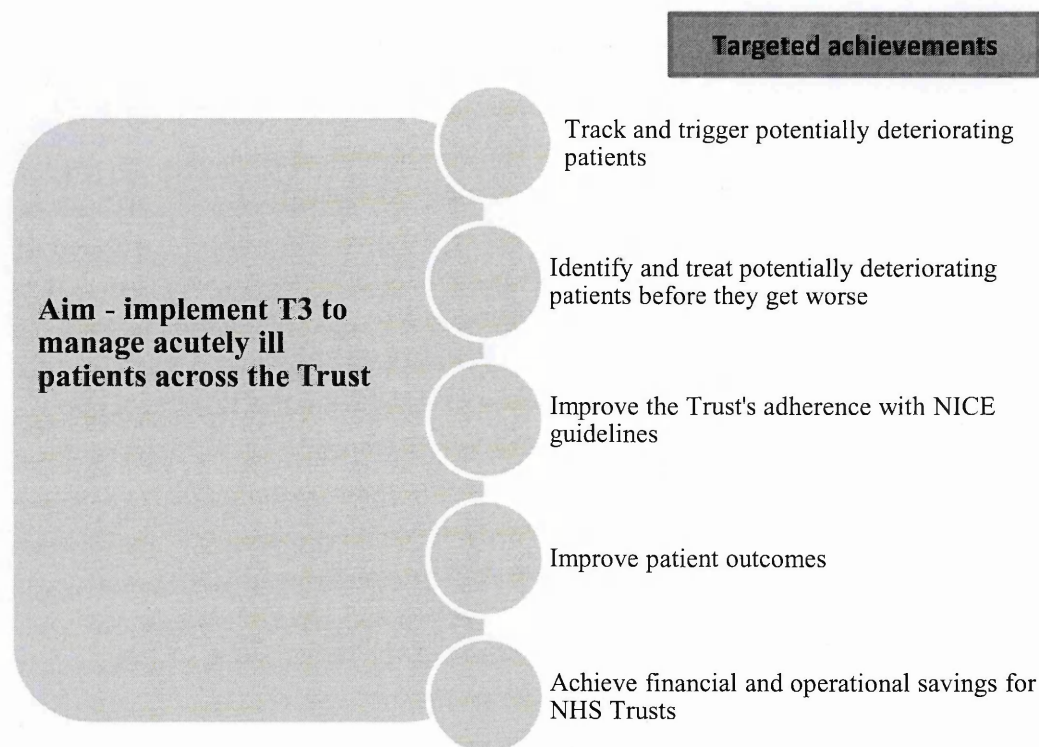


Figure 7.1 Problems that were expected to be solved by implementing and using T3

#### 7.4.2 How T3 works

T3 was an expert software CDSS that was accessed via a handheld personal digital assistant (PDA). It was designed for use by nurses and healthcare workers at the patients' bedside. It was adopted by the study Trust in 2007 to replace a paper based MEWS chart whereby nurses and other healthcare workers used to input clinical information then calculate MEWS scores manually using parameters printed at the back of the chart. T3 had 5 core applications designed to capture and analyse clinical data at the point of care as well as providing additional functions for data mining and management reporting (see Figure 7.2). The Patient Flow, Nurse and Clinical applications were the most widely used in the study Trust to record and view clinical observations.



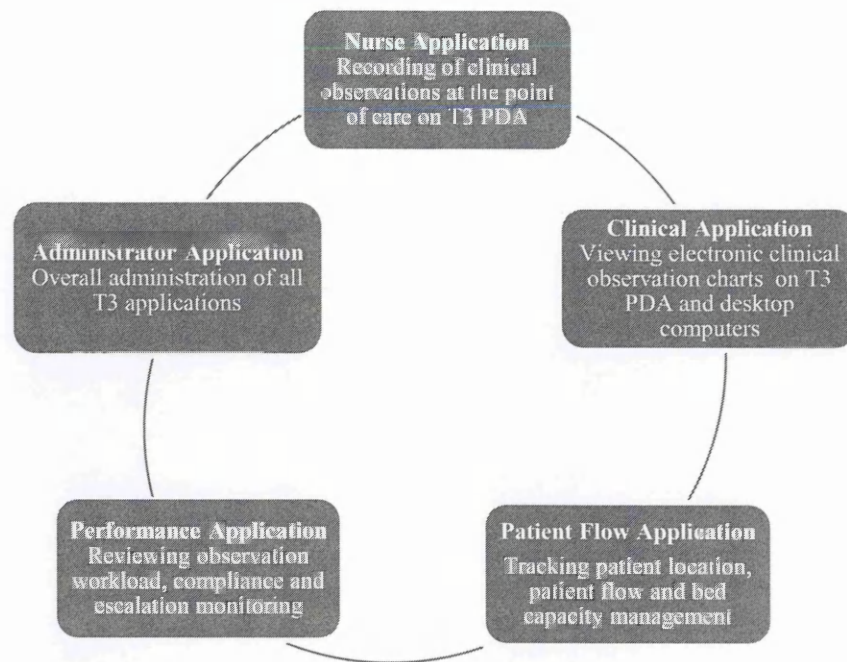


Figure 7.2 T3 core applications

The original T3 design premise was that nurses and healthcare care assistants would input the patient's blood pressure, temperature, respiratory rate and heart rate into T3's Nurse Application. A date and time stamp would be given for every set of clinical observations recorded. T3 then analysed the entered data against set parameters based on the MEWS protocol and gave an early warning score on its screen. It also suggested the frequency of subsequent observations accordingly. Where observations were outside normal ranges, T3 would advise the user to inform a senior nurse or doctor depending on the severity of the abnormality. Additionally, it also had the capability to prompt users to start interventions such as oxygen therapy and secure the timely presence of appropriately skilled clinical help for those patients exhibiting signs of physiological deterioration. T3's inbuilt prompts, alerts and reminders were designed to reduce the risk of missing key observations or identifying and responding to patient deterioration. The alert levels and parameters were based on NICE guideline 50 and NPSA recommendations on how acutely

ill patients should be managed in acute hospital environments. Clinical observation trends could be viewed on the T3 PDA as well as desktop computers within the study Trust's network. During ward rounds, the medical teams used the study Trust's "computers on wheels" to view clinical observations and trends while moving from patient to patient. T3 was linked by a wireless local area network to the Trust's electronic clinical results reporting system. This link allowed accessibility to patients' clinical information to members of the multi-disciplinary teams from dispersed locations across the hospital (see Figure 7.3). Viewing trends was important because it allowed healthcare workers to "track" the patient's clinical vital signs over a period of time.

The Critical Care Outreach team remotely monitored MEWS scores, and offered telephone support and prompts for wards to commence interventions where indicated. Where necessary, the Outreach team physically went to ward areas to review patients and liaise directly with the responsible medical and nursing teams. Sometimes they initiated the appropriate clinical interventions to prevent further patient deterioration and avoid admission to the intensive care unit. T3 also gave an automatic visual alert when clinical observations were overdue by displaying a red icon next to the patient's name. The Nurse Application was linked directly to the Patient Flow Application, which provided a prompt for users to enter the ward name, patient's bay and bed location when the first sets of clinical observations were entered. Users were required to update this information whenever a patient was transferred to another bed or different ward. The Clinical Application displayed electronic patient observation charts through T3 PDAs, desktop computers and the "computers on wheels". Patient lists could be viewed by ward, consultant, MEWS scores or individual searches using NHS or locally allocated hospital

numbers. The Performance Application was designed to give ward managers, hospital matrons and senior clinicians reports about how their individual clinical areas were performing. They were then expected to use these reports to improve performance and patient outcomes in their clinical areas. The Administrator Application was used by the study Trust's ICT department for the overall management of all T3 applications.

### **7.4.3 Expected T3 benefits**

There were many expected benefits for the various actors involved with T3. For patients, the main benefit was that they would receive improved care that was based on national guidelines as set out by NICE. Automated MEWS calculation was expected to improve patient safety through prompt recognition of physiological deterioration and timely instigation of interventions. For nurses and healthcare assistants, T3 provided a quicker, more reliable and safer way to calculate MEWS and track patients. The additional triggers provided through prompts, alerts and reminders were designed to upskill junior and inexperienced nurses to provide care to acutely ill patients. Other benefits for nurses included easy clinical data accessibility and visual presentation to highlight key aspects of clinical observations and a move away from variable and often illegible handwritten paper-based forms. Nurses and healthcare assistants were expected to save time and to focus more on attending to patients' needs rather than manually calculating MEWS and writing them in paper charts. The main benefit for the Critical Care Outreach team was that T3 would enable them to view all in-patients from their remote base. This facility was expected to help to prioritise their workload, identify and offer support to non-compliant wards, identify learning needs and tailor training accordingly. The Outreach team's PDAs also provided alerts of patients that had a MEWS score of four and above. This additional

function was expected to help the team to identify patients that needed to be reviewed and to prioritise the commencement of interventions. For doctors, the main T3 benefit was that they would prioritise their workload by evaluating and responding to patients that were most in need based on their MEWS scores. Additionally, they could also provide advice to wards and initiate clinical management plans over the phone from remote locations. For clinical and service managers, T3's patient tracking and triggering functions provided an auditable and guideline compliant way to monitor patients transferred across the hospital and report the hospital's performance. Based on the Performance Application's management reports, they could also benchmark their performance against other comparable NHS Trusts, identify failing wards and implement remedial actions. Additionally, they could also evaluate hospital bed utilisation to ensure that patients were in the most appropriate specialties depending on their condition and expedite discharge planning. For the Trust Board and departmental Clinical Directors, T3 was a safe and validated product that would help them to meet the study Trust's organisational duty of care by meeting NICE guidelines and NPSA recommendations and satisfying key stakeholders such as the government, commissioners and patients. As such, they could justify resource allocation for T3 procurement. Additionally, because T3 was provided by a commercial third party supplier, the study Trust expected continued technical support and capability to allow revisions and additions to T3 software algorithms in line with guideline changes and advances in clinical practice.

#### **7.4.4 T3 implementation in the study Trust**

T3 was implemented in the study Trust at the beginning of 2007, starting with a single ward pilot, then Trust wide rollout. T3LeadConsultant revealed that T3 was found to be

“unusable” at the time of adoption. This conclusion was based on informal assessments that were carried out by the implementation team at the initial stages of the pilot implementation. The study Trust worked with the T3 supplier to establish the clinical needs and map T3 to the various clinical processes before Trust-wide rollout. There was also no business continuity plan at this stage. Consequently, T3LeadConsultant co-developed a business continuity plan with the supplier to ensure continued post implementation support. Additionally, a T3 operational policy was developed for use by clinical staff with responsibilities for recording or reviewing patients’ clinical observations. This was done to ensure the safe use of T3 applications in line with the study Trust’s corporate objectives. The T3 operational policy also stipulated the duties and responsibilities of the key actors involved in the implementation and use of T3 (see Table 7.1). The policy could be updated whenever there are changes to guidelines, service requirements, processes, system/software upgrades and new system module releases. The initial plan was to have tablet computers by every bedside but due to high set up costs, the study Trust settled for a few handheld PDAs on each ward.

<b>Actors</b>	<b>Duties and responsibilities</b>
<b>T3 Steering Group (Trust board members, senior clinicians, and senior nurse managers)</b>	<ul style="list-style-type: none"> <li>• Support the implementation of T3 in clinical areas, its updates and any related systems in line with the Trust's corporate objectives</li> <li>• Ensure appropriate use of T3, its associated modules and any future developments</li> <li>• Ensure T3 supports clinical systems, improves patient safety and that it is fully integrated into the Trust's clinical and operational processes</li> <li>• Identify and ensure the realisation of clinical and operational benefits that can be achieved from T3 implementation</li> </ul>
<b>T3 Clinical Users (Nurses, doctors and allied healthcare workers)</b>	<ul style="list-style-type: none"> <li>• Safe and correct use of T3 in accordance with the Trust's operational policy</li> <li>• Adherence with published national guidelines and local clinical protocols</li> <li>• Appropriately escalate care within existing multidisciplinary structures in respective clinical departments and the Critical Care Outreach team</li> </ul>
<b>Clinical line managers (modern matrons and ward managers)</b>	<ul style="list-style-type: none"> <li>• Implement appropriate measures within respective clinical departments to support T3 use</li> <li>• Ensure adherence with T3 operational policy</li> <li>• Ensure adherence with published national guidelines, Trust clinical protocols and departmental objectives</li> </ul>
<b>ICT department</b>	<ul style="list-style-type: none"> <li>• Manage operational and technical issues related to T3 software and hardware</li> <li>• Manage service user alerts and system availability</li> <li>• Implement new functions and updates</li> <li>• Provide training for new and existing users for software and hardware devices</li> </ul>
<b>T3 Supplier</b>	<ul style="list-style-type: none"> <li>• Provide technical support to maintain reliable operation of T3</li> <li>• Offer appropriate training for the safe use T3</li> <li>• Respond to and resolve support calls for software and hardware</li> </ul>

Table 7.1 Duties and responsibilities of key actors involved with T3

### **T3 pilot implementation**

In January 2007, T3 pilot implementation was commenced on the study Trust's renal ward. This ward was selected for the pilot because it had high patient turnover and adequate medical cover to support the nursing staff during implementation. T3LeadNurse, who had previously worked as a senior nurse manager and had been involved with the implementation of clinical software, headed the T3 pilot implementation team. Other team members included T3LeadConsultant, T3ImplementationNurse, ICT department project managers and technicians, renal consultants, nurse managers and nursing teams and T3 company representatives. The implementation team focused on the integration of T3 PDAs onto the ward's clinical processes and with existing electronic systems. User training and support was provided on a daily basis. T3ImplementationNurse noted that the implementation team also had to set up user accounts for a significant number of nurses and healthcare support workers who did not have access to the study Trust's Intranet at the time. This was unexpected because the majority of nurses and healthcare workers were expected to be already using the Trust intranet to access their emails and to view test results among other things.

It was a big job [the implementation]... quite a lot of nurses didn't even have computer login details or access to the system for electronic charts...lots of facilitation, getting line managers to sign things off, just trying to make things easy for them... account set up [problems with] were unexpected , although there was an inkling... little has changed regardless of the number of projects that have been rolled out since... we are still bringing forms to get people logged in... it's becoming role specific, nurses have access, but there is a huge gap with healthcare

assistants, but now with ESR, it's becoming more important for them to use computers...

#### T3LeadNurse

At the end of each day, the implementation team met to discuss and resolve any issues arising. For the first 6 months, clinical observations were recorded on both T3 and the paper-based MEWS charts. For the latter, the MEWS scores were manually calculated. The dual recording of the clinical observations was reported to be very labour intensive. During this period, changes were made to the T3 software algorithms to adequately track patient movement and instances where observations were delayed or missed. The implementation team reported significant resistance from nursing and medical teams who perceived T3 as an additional and irrelevant system. The T3 PDA also encountered various technical problems such as delays in loading data, failures in synching with the electronic clinical results reporting system, intermittent data losses, low battery life, and poor presentation of clinical trends. T3LeadNurse noted that most nurses and clinicians preferred to use paper-based observation charts because they clearly showed clinical trends between sets of clinical observations and patient progress over time. There were also problems with accessing data from the PDAs due to the connectivity issues caused by layout of the new hospital.

The first PDA broke regularly, batteries were poor and the screen dented... in-between patients, it would say "we are networking and buffering", and only allowed you so many observations before the battery ran out... the stylus was not



attached and people were using ball pens... ink all over the screen, got stuck in itself...

#### T3LeadNurse

The implementation team spent several weeks identifying 'blind spots', while the T3 suppliers worked on optimising wireless connectivity. The renal ward pilot was extended from 6 to 9 months to allow the T3 implementation team and suppliers time to resolve outstanding issues. The layout of T3 results was changed to allow ease of use and integration with the existing clinical results reporting system. Paper-based charts were completely withdrawn from the renal ward at the end of the extended pilot in October 2007.

#### **Trust-wide T3 implementation**

In September 2007, a project plan was drawn up by the T3 implementation team leaders for Trust-wide T3 rollout. The implementation team was expanded to include 6 senior nurses split into 2 teams of 3. These nurses had extensive experience in different clinical specialties, patient-flow management and training and development. They were responsible for training and supporting ward teams as well as troubleshooting any issues arising during and following the implementation. T3 went live 2 wards at a time, with a small number of wards allocated per team. The plan was to rollout T3 rapidly, using the SCOPE project management model and involving the key actors such as nurses, healthcare assistants, doctors and line managers. Rollout was based on the wards' geographical location, clinical specialties, and interrelationships and co-dependencies between ward areas to ensure workflow continuity. T3LeadNurse noted that the implementation was flexible, in order to facilitate the necessary changes and respond to issues arising in ward

areas. The implementation team used a checklist to systematically introduce T3, starting with the ward staff on duty list, scheduling training, establishing the number of PDAs that were required and where they would be located on the ward. Once a group of wards was identified, the implementation team had a week of preparation with ward managers. This was followed by 2 weeks of “intense training” (T3LeadNurse) using T3 and paper-based MEWS charts, then a paperless period of 2 weeks. On completion of the 2 week paperless period, an agreement was signed with respective ward manager to acknowledge that they were satisfied that the ward had been adequately prepared for transition to paperless clinical observations. During the second week, the implementation team would start preparing for the next group of wards, while allowing time for additional support and ensuring that all staff had been adequately trained.

Our approach to user training was flexible and pretty mixed... It was easier to fit training around the nurses’ station rather than sisters’ office, so the nurses were still accessible on the ward if [they were] needed. It was broken training, because emergency buzzers, phones ringing... but they [the nurses] had faith in that we would let them attend to their work, and they actually concentrated during training... We also went on observation rounds, and were all in nurses’ uniforms. We were able to freely go around the ward, be with patients and looked professional... it helps when patients know that you are a nurse...

T3LeadNurse

The implementation team made various changes to their rollout plan to suit different working patterns across the study Trust and to resolve logistical challenges during both implementation phases. T3LeadNurse noted that in some areas, T3 was deemed

unsuitable for implementation and collaborative work was undertaken with the relevant clinical departments to revise the software algorithms to local needs and adapt it to suit the processes.

If you walk on to a ward and realise there is no chance [i.e., if wards are busy]... it's being responsive, having a backup ward, flexibility to move between wards and being known around the Trust gave it some credibility... you have the contextual piece, not somebody who doesn't have a clue about doing clinical observations... that was important... bringing feedback and responding to things, removing blockages, resolving problems... encouraging culture change to an electronic process...

T3LeadNurse

Some clinical specialties did not believe that T3 would be beneficial to their wards. For example, one of the line managers from a specialist ward (WardSister1) questioned the suitability of T3 for such clinical areas, noting that their existing monitoring systems had more advanced capabilities than T3. She particularly questioned the study Trust's motives for introducing T3 across the hospital, noting that it was more likely to cover the Trust in case of litigation rather than improving patient outcomes. However, she also acknowledged T3's relevance in general wards and for inexperienced nurses.

It's a bit like big brother watching you... to keep an eye on people who don't know what they are doing, in a helpful way. It's like an idiot's guide when to do observations. It takes away professional judgment and tells you what and when to do it... It's for the Trust and everyone must have it... legally they [senior Trust managers] can prove that patients' observations have been done because they

have it down on a computerised system. I am not dead against it... we got used to it and I probably wouldn't want to go back to paper charts now...The idea was that doctors aren't called unnecessarily, and they can look up for observations without having to go to the ward. We do that ourselves when we get referrals from other wards [general wards]...

WardSister1

By the end of 2013, T3 had been implemented in all adult in-patient beds except for the Accident and Emergency, maternity services, theatres and recovery, and general critical care departments. The T3 implementation team worked closely with the supplier throughout the implementation process. T3LeadNurse noted that they received extensive support from T3 developers throughout both implementation phases. This included project management and T3 customisation to suit local needs where feasible. T3 modules were also updated and sometimes redeveloped to suit clinical processes, especially in specialist wards.

It was an agile approach to software development, going back and forth and trying to support the rollout. It was their [T3 supplier] first ever roll out of such magnitude so they were quite willing to work with us and provide the functionalities that were needed... It had benefits for them because if we were asking for something, other Trusts were bound to want the same... they actually began to have a comprehensive package that they could sell . . . they wanted us

to roll it out Trust wide and could then say, “we have this system, which we rolled out in this big university hospital in 6 months...”

T3LeadNurse

On completion of the Trustwide implementation, T3LeadNurse and T3ImplementationNurse were seconded to the ICT department. This was done to allow continuity and ongoing support for the wards, as well as to ensure that T3 was being used appropriately and implementing contingency plans where necessary. The study Trust’s ICT department and the supplier jointly managed T3 maintenance. Updated T3 PDAs were introduced following Trust wide implementation. This was done after a trial involving nurses and the infection control department.

We are out there... it is involving everybody and finding out requirements... we were not brought in for the initial T3 implementation, so I don’t know how this was addressed but we are certainly involving all user groups... you can’t do something this big without involving the users...

T3ImplementationNurse

Although, there was no documentation of this trial. The new PDAs were reported to be a vast improvement on the old ones. However, there were some changes to how T3 was originally intended to be used. The original design premise was that alerts would automatically be sent to the Critical Care Outreach team once MEWS scores of 4 and above were recorded. This led to large numbers of alerts which instead of helping became a hindrance to the Outreach team. Most were false alerts and in the majority of cases, the necessary interventions would already have been started. This led to the alert function

being switched off by the Outreach team. Instead, they logged into T3 early in the morning when they got handover from the Hospital at Night team to review MEWS scores of 4 and above. They then checked again at lunchtime and responded to direct referrals from wards and medical teams in-between. All the work undertaken by the Outreach team and their communication with ward areas and clinical teams was documented in patients' paper-based medical notes and added to the MedICUs national database. This information was used for audit purposes and as a reference point for future admissions and interventions. Ward areas adopted various methods to ensure that observations were appropriately recorded and followed up. It was noted that most nurses continued to record observations on the discontinued paper charts, handover sheets and also in patient clinical records. Some were reported to be scribbling observations on pieces of paper and then loading them onto T3 afterwards. This posed risks of data loss, incorrect submissions as well as having the wrong date and time stamp.

The implementation team also reported significant resistance from doctors, especially those from surgical specialties during implementation. T3LeadNurse noted that most doctors preferred the paper-based MEWS charts because they could see the trends better than on T3. In some cases T3LeadNurse revealed that she had to request T3LeadConsultant to "come along to meetings" with senior surgeons to explain why they had to implement T3. She noted that T3LeadConsultant "would then say exactly the same things as I would have said" and all the senior clinicians would agree with him (T3LeadConsultant) because they were "big customers" of the critical care department's services and as such they wanted to be seen to be complying to ensure that they secured critical care beds for their patients. It would appear that T3 implementation was being

used as a political tool to enforce compliance and also allocate limited resources to those who came on board. However, there was ultimately poor “buy in” by key stakeholders, which goes against the core principles of the benefits realisation model.

There was more reluctance from medical staff. Since it was a big change for nurses rather than doctors, you would have expected it to be the other way round... some doctors insisted that nurses print all observations, some wanted charts at the end of the bed, which defeats the object of T3... I admire nurses, we gave them something, not necessarily the best, and they learnt and make their own adaptation...the new PDAs are better, so its breaking the bad habits formed out of necessity [with the old PDA], that's how they did it because they had to... our challenge is trying to prove to them the system is now reliable and fits the purpose...

#### T3ImplementationNurse

WardSister1 argued that despite some advantages, T3 could not replace experienced nurses and could potentially deskill ward nurses. She also argued that resources could be better spent on training nurses and ensuring that wards were adequately staffed and the appropriate skill mix maintained.

Night sisters and the Outreach team can now identify them [deteriorating patients] and intervene appropriately... for a system that I imagine costs a lot of money, I'm not sure whether it's worth it. If it has cost more than making sure that you have enough staff on the wards, senior nurses to keep an eye on things rather than machines on the ward ... you can't beat someone being physically there...it is

taking away people's ability to make decisions and in the long term may deskill people...

WardSister1

## 7.5 T3 evaluations

Prior to adoption by the study Trust, T3 developers and the developing Trust performed various evaluations. These evaluations included clinical validation of the MEWS protocol (supporting T3), evaluation of the long-term effects of MEWS scores in acute medical settings and comparison between manual paper-based MEWS calculations and those electronically calculated using T3. Additionally, T3 was validated to assess its effectiveness in the surveillance and management of acutely ill patients. These studies concluded that the MEWS was a useful tool to identify and manage deteriorating patients in acute settings. Manual MEWS calculations were found to be inconsistent, often with omissions and incorrect entries. In contrast, T3 was reported to be faster, safer and more accurate. Significant reductions in ITU admissions and deaths were also noted in the developing Trust. Table 7.2 shows the evaluations that were carried out within the study Trust following T3's implementation in chronological order.



Method	Timing and duration	Evaluators	Evaluation purposes	Target audiences	Results	Uses of evaluations
Informal observations and assessments during the pilot and Trust-wide implementation	During the 9 month pilot and over one year during the Trust-wide implementation	T3 implementation team and line managers	To ensure full implementation of T3 and resolve issues arising	T3 Steering Committee and Trust Board	Issues arising resolved in the course of implementation	To demonstrate that the system has been implemented across the organisation
Benefits realisation studies	12 months and 7 years post implementation	Project Lead Consultant and Outreach Lead Nurse	Assess whether expected benefits had been realised	T3 Steering Committee, Trust Board and Commissioners	T3 found to have helped to overwhelmingly improve and maintain all key indicators	To demonstrate that the planned/expected benefits have been achieved in order to provide a case for continuing funding
IQNARC audits	Performed monthly	Critical Care Department and the IQNARC Case Mix Team	To assess the effectiveness of the Critical Care service	Trust Clinical Governance team, Trust Board, Commissioners and regulators	The Trust has achieved and maintained all key performance indicators	Use the data from the CDSS to argue for continued resourcing of an operational unit.
Dr Foster audits	Data submitted monthly by the	Critical Care Department and Dr Foster	To monitor the effectiveness of acute care services	Trust Clinical Governance team, Trust Board,	Monitoring and reporting of deaths improved. Trust	Use the implementation of the CDSS to

	Critical Care Department		and deaths rates in the Trust	Commissioners and regulators	now reporting fewer deaths and acute services found to be adequate	underpin the argument that certain quality procedures were place.
Outreach service evaluation	Performed in 2012 over two months	Critical Care Outreach Team	To evaluate the uptake and impact of the Outreach service	Critical Care Department	T3 found to have improved the Outreach team's efficiency and patient outcomes	Use the data from the CDSS to argue for continued resourcing of an operational unit.
Root cause analyses	Performed when adverse events are reported	T3 Lead Consultant	To assess the cause of adverse events in hospital and put in place remedial actions	Clinicians, Trust Clinical Governance team, Trust Board and regulators	Outcomes of investigations shared with departments, shared learning and identification and improvement of systemic failures	Root cause analysis now routinely used by all clinical departments with support from the Critical Care Department
Informal evaluations	Pre and post implementation	The implementation team and T3 users	To assess the effects of T3 on clinical processes and existing systems	T3 users, line managers, implementation team, T3 steering Committee and Trust Board	Improvements noted in the use of T3 reported.	Appropriate changes made to resolve operational and systemic problems noted

Table 7.2 Evaluations undertaken in the study Trust (in chronological order)

### 7.5.1 Evaluations during implementation

From 2007 to 2009, a multidisciplinary team of nurses, doctors, project managers, IT technicians and the T3 supplier carried out various evaluations during its implementation in the study Trust. The purposes of these evaluations were to review the implementation process, assess T3's suitability to the study Trust's clinical processes and workflow and to obtain feedback from wards to inform the development team. Additionally, the implementation team sought to evaluate the effectiveness of user training delivered, deployment and testing of PDAs, checking integration with legacy systems and the transition from paper-based to electronic clinical observations.

T3LeadNurse noted that the majority of these evaluations were informal and undocumented. They primarily focused on responding to daily implementation activities through multi-disciplinary 'wash up' meetings to address any issues arising. The team performed audits of ward profiles and used standard checklists to establish hardware needs in ward areas, staff training needs and work out contingency plans. Following implementation on each ward, the checklists were completed and signed by T3LeadNurse and respective ward managers to confirm that they understood the T3 operational policy and business continuity plan as well as their ward's readiness to go paperless. T3LeadNurse noted that the "culture" around observation taking and recording and the use of the MEWS had improved across the study Trust. However, she also highlighted that it was difficult to adapt and link T3 with the Trust's existing clinical results reporting system and clinical processes. This was mainly related to multiple problems with the T3 PDA such as poor battery life, poor layout of clinical trends and connectivity with the study Trust's Intranet, persistent data loss and a detached stylus. T3LeadNurse argued that some of these problems could have been identified through pre-

implementation evaluation of T3's suitability to the study Trust. The implementation team also reported that "buy in" by clinical users remained low throughout the implementation phases. They noted that most nurses did not trust T3 due to general apathy towards technology and its associated technical problems. As such, they continued to use paper charts alongside T3 as well as writing in patients' clinical notes. In some specialist clinical areas, it was found that T3 was unsuitable for the patients' clinical needs. In response, T3 implementation was delayed to allow changes to the software algorithms and addition of new functions. T3LeadNurse noted that the T3 suppliers were supportive and worked with the implementation team to respond to user feedback and make the necessary software and hardware adaptations as required at the time.

I don't think we were completely aware of the challenges [with T3 implementation]...

I wasn't involved in evaluating it [T3] prior to purchase, but on first impressions seemed it would do as expected; it records the observations, it calculates MEWS, but when issues arose... it's about making changes to make it a better system, improve things for users, changes to oxygen screen, patient flow element is completely new... looks nothing like what they originally presented to us. Still work in progress... they were very open at that point in time to make changes required regarding patient safety and clarity. At 3pm every day, the team would meet and bring back issues, deal with them quickly, being responsive and listening to people, taking on board their ideas...

T3LeadNurse

The unavailability of T3 implementation and evaluation documentation made it difficult to grasp the full breath of evaluations carried out and corresponding results. It would appear

that the implementation team mostly responded to activities rather than following a clearly structured evaluation format using a clear methodology. Instead, professional judgements by the individual actors and clinical groups, particularly members of the critical care department were taken as evaluations. There was also no formal implementation report to corroborate the verbal accounts of various actors involved in the implementation and evaluation of T3.

### **7.5.2 Benefits realisation studies**

In 2008, T3LeadConsultant carried out a “formal benefits realisation study” with support from the Critical Care Outreach Team Lead Nurse. The purpose of the study was to assess one year following implementation whether T3 had achieved the expected operational, financial and clinical benefits as those achieved by the developing Trust. These benefits included critical care avoidance rates, reduction of cardiac arrests and length of stay in intensive care units, compliance with NICE guideline 50, uptake of the critical care outreach service, releasing nurses’ time, tracking patients and identifying patients fit for discharge to ensure appropriate use of resources. The benefits realisation study was repeated at the end of 2013. The second study was triggered by the tender process to replace or continue to use T3 by establishing whether the improvements identified in the initial study had been sustained. This study was performed by the T3LeadConsultant and the Critical Care Outreach Lead Nurse. It also included some key issues that were national priorities for the NHS at the time. These included compliance with the Sepsis 6 bundle for treatment of patients with severe infections, audit of the Dr Foster hospital standardised mortality ratio, preventing, tracking and managing infection outbreaks, assessment of quality of life following critical care discharge, compliance with the national emergency pathway, assessing the effectiveness of the critical care outreach service, hospital and user performance and improvement in crude mortality rates.

T3LeadConsultant argued that T3 was continually being evaluated and improved since it was adopted to show clearly the benefits such as improving operational processes and ultimately patient outcomes and to justify the study Trust's investment in T3. Both T3LeadConsultant and T3LeadNurse reported that T3 benefits had been "overwhelmingly proven" on all the outcomes that had been measured. T3LeadConsultant argued that justifying the Trust's investment in T3 was important because the like many NHS Trusts at the time, it was facing significant financial challenges.

The current cost of replacing T3 is about £450 000 annually so the Trust has to justify the IT investment. Like many Trusts, we have a significant deficit and have to introduce cost improvements, which is often quite difficult. It was therefore imperative to analyse the benefits [of T3], and continually improve this product, especially towards the mobile platform, which is now cheaper and more readily accessible... the additional functions we would want from them [T3 supplier] are too expensive... Existing functions such as Performance Application could be better utilised by nurse managers and departmental managers...

T3LeadConsultant

It was difficult to assess the effectiveness of the benefits realisation studies due to lack of evaluation documentation. Although the study reports and related presentations that had been delivered to various stakeholders within and outside the Trust, peer groups, potential users and national platforms were said to be available, the researcher was not given access to them because of the ongoing tender process. Additionally, T3LeadConsultant revealed that he could not "release any information" because he was looking to publish a series of research papers related to the adoption and evaluation of T3 in collaboration with the developing Trust

and T3 suppliers. It was unclear whether the benefits realisation model had been used. However, the lack of involvement of other actors outside the critical care department suggests that the studies were primarily focused on departmental aspects and how they relate to the wider Trust rather than Trustwide issues. There was also no apparent awareness of any shortcomings or “blind spots” in both benefits evaluation studies. It appeared the evaluators were confident of the robustness of the methodologies they had applied and the results that had been produced.

### **7.5.3 IQNARC audits**

Since the late 1990s, the study Trust had been submitting performance data to ICNARC in line with national audit requirements. The mandatory key performance indicators of the ICNARC audits included mortality, management of sepsis and length of stay in critical care. These indicators were driven by issues that were considered important at national level. As part of ongoing service evaluation, the Critical Care Outreach team submitted monthly data regarding all their activities to ICNARC’s Case Mix Programme through the MediCUs database. The MediCUs software was developed by a private company to manage and monitor the performance of all NHS critical care outreach services. The submitted data were then analysed and results published by IQNARC on its national database. Various management reports were generated periodically to help critical care departments to assess the effectiveness of care delivered. The MediCUs software provided a database resource to collect and organise information for patients who were seen by the Outreach Service. It also provided tools to analyse the effectiveness of interventions delivered and generated annual activity reports on a wide range of issues such as comparisons of MEWS, effectiveness of interventions that were carried out, escalation decisions made, management at ward level, Do Not Attempt

Resuscitation orders initiated, end of life decisions made, and whether critical care admissions had been avoided. The critical care department carried out quarterly performance reviews and benchmarked itself against other NHS Trusts of similar size using data available on the national database.

T3LeadConsultant noted that following T3's implementation, the study Trust had improved compliance with key indicators for ICNARC's Case Mix Programme from 94% to 100%. Internal audits that were performed by the emergency department since 2010 showed that deaths following sepsis were lower than the national case mix data sets. Additional audits undertaken internally to measure outcomes of patients who had suffered cardiac arrest in hospital had been in line with national averages reported by the Resuscitation Council (UK). At the end of 2014, the Trust was preparing to submit the required minimum data sets for ICNARC's Cardiac Arrest Audit. T3LeadNurse also noted that admissions to critical care had been reduced because the Outreach team were identifying and responding to deteriorating patients quicker. However, she also acknowledged that these improvements were "coincidental" and could not be solely attributed to T3.

There was a coincidental reduction in key performance indicators using evidence from ICNARC data, although not solely attributable to T3... for example, length of stay in critical care and admissions averted from critical care because the Outreach team are getting to patients earlier and escalating cases. Proven benefits such as ICNARC data and reduced length of stay, it's an evaluation... describing the benefits going forward



for the business case to support the new tender... asking questions such as 'are we continuing to sustain those lower mortality rates from 6 years ago?'

T3LeadNurse

The reported improvements in ICNARC audits were focused on general patient outcomes as defined at national level and did not directly evaluate T3. Despite the general consensus within the critical care department that improvements in key performance indicators resulted from T3, there was no clear causation. However, it was possible that the accessibility of patient information through both T3 and MedICUs could have improved the reporting of key performance indicators. Despite the lack of direct cause and effect, these reported benefits were published in the study Trust's annual accounts and reports, as well as supporting the developments of the business case towards the continued use of T3 in the study Trust.

#### **7.5.4 Dr Foster mortality audits**

Like many NHS Trusts, the study Trust used Dr Foster's Real Time Monitoring tool to monitor its clinical outcomes. A key element of this process was the Dr Foster hospital standardised mortality ratio (HSMR) audit. Data collection is based on the Dr Foster alert system to monitor specific diagnoses that are known to cause death and procedure related deaths in NHS hospitals. In 2009, Dr Foster Intelligence rated the study Trust as Level 1, with mortality rates of 114.2 annually, which were the third highest in the region. However, the study Trust challenged these results by issuing a press release which highlighted various initiatives which they had implemented, favourable ratings from other regulatory bodies and awards such as the ICT Excellence Award for best Added Value Project (T3). They argued that Dr Foster's results, which were based on a questionnaire survey and other submitted data lacked depth

and failed to recognise the high quality of care that had been achieved by the study Trust, ratings by other regulatory bodies that the Trust considered more in-depth and independent reviews. Additionally they noted that the methodology used for analysis had not been shared with the Trust and had been based on incomplete data. The purpose of this press release was to reassure patients and other stakeholders of the high levels of patient safety maintained by the study Trust by highlighting the inaccuracies in the Dr Foster rating. Around the same time, a study by Mohammed et al. [2009] concluded that Dr Foster Intelligence's case mix methodology was unsafe for case mix adjustment and was likely to increase the bias which the case mix was intended to reduce and lacked credibility. Following the study Trust's press release, the regional Strategic Health Authority tasked Dr Foster Intelligence to investigate the causes of high mortality and work together with the study Trust to make the necessary improvements. Dr Foster's investigation found the study Trust's monitoring and reporting processes to be reactive and poorly understood by clinicians and managers. Working with the study Trust's clinical governance team and senior clinicians, Dr Foster integrated its Real Time monitoring tool into the study Trust's existing systems. They recommended that mortality data should be monitored at both whole Trust and specialty levels using Dr Foster tools. Additionally, the study Trust's Quality Governance Group was tasked with reporting the number and causes of deaths quarterly to the Patient Safety and Mortality Review Committees. All deaths were reviewed in departmental meetings to monitor clinical performance, share learning and highlight and respond to systemic concerns identified.

Since 2010, the study Trust's HMSR dropped below 100 and its ranking in the region had significantly improved by 2014. These results were reported in the study Trust's quality accounts and annual reports and were attributed to the use of T3 and subsequent

improvements in work processes. However, the impact of T3 on HMSR monitoring and reporting had not been evaluated and thus no direct cause and effect link could be established. In fact, by the time of the controversial HMSR rating, T3 had already been in use for nearly 2 years in the study Trust. It was more likely that the improvements reported were a result of changes to reporting deaths following direct intervention and continuing support from Dr Foster. Also challenging the Dr Foster rating goes to show the difficulties in measuring performance where multiple initiatives are concurrently running. However, the same principle was not applied when directly linking T3 to the dramatic improvements in HMSR improvements. In 2014, the Dr Foster Intelligence methodology was also challenged by leading academic, Professor Nick Black (Triggle, 2014), who argued that the key measurements for hospital death rates were misleading and should be ignored.

I don't think there's any value in the publication of HSMR and I'd go further, I think it's actually a distraction because it gives... a misleading idea of the quality of care of a hospital... Personally, I would suggest that the public ignore them...

Professor Black (Triggle, 2014)

In response, Dr Foster Intelligence argued that while the HMSR data were not necessarily confirmation of poor care, they were important indicators of potential problems. They noted the example of death rates which had been used to identify problems such as those which later unfolded at Mid Staffordshire hospital.

### **7.5.5 Critical Care Outreach service evaluation**

In 2012, the Critical Care Outreach team carried out an evaluation of its services. The purpose of the evaluation was to assess the utilisation of the outreach service and its impact across

the hospital. OutreachSister2 also noted that the evaluation was also an important way to “justify our existence” and to remind decision makers of “our worth” to the Trust. Proving their worth was important because it would ensure their continued existence, ring-fencing of available funds and potentially attract more funding to expand their service. The evaluation report was mainly descriptive, noting the key achievements of the Outreach service since 2006. The report concluded that the Outreach service had vastly improved since implementation of T3. Improvements highlighted included increased efficiency of the outreach service, accessibility of real time patient data and quicker response to potentially deteriorating patients using both referrals from wards and remote access monitoring. Consequently the necessary interventions were being instigated timeously and patients were now being diverted from intensive care. OutreachSister2 noted that there had been “huge changes” in ward nurses’ skills and practice and improved identification of “problem areas”. The Outreach team reported that routine audits that used to take several days of manual work were now being done in “just a few clicks”. Examples of these audits included the previously paper-based tracheostomy audit of patients admitted to the Critical Care Department.

Anecdotally, from the beginning looking at the skills of the nurses, there is a huge change in practice... people are more proactive now, early interventions started, which never used to happen... we also teach on the high dependency course, assessment of patients... this empowers the nurses to think and links with SBAR, ensuring better

communication... fits with NICE guideline 50, wasn't formally assessed before, now it will be a formal course, management of the acutely ill...

T3OutreachSister1

Additionally, the availability of information such as real time reports of patients with high MEWS, ward performance, clinical specialties, ward or consultant enabled the team to identify and troubleshoot issues arising. They also reported improvements in staffing and skill mix on the wards, establishment of training needs and implementation of relevant training programmes for nurses and healthcare assistants. In 2013, the Outreach team started to assess nurses in various wards on the skills for high dependency course. The course involved clinical teaching to enable nurses to systematically assess patients on wards and improve communication when escalating unstable and deteriorating patients to doctors and the Outreach team. It linked with the MEWS protocol that was integrated into T3. The course was supported by line managers and senior clinicians. Further improvements were expected when additional software updates such as the sepsis pathway were implemented.

The Outreach team noted that the effectiveness of training delivered was difficult to measure because previously nurses' clinical skills were not formally assessed. However, they argued that despite the lack of evaluation, there were improvements in awareness and response to patient deterioration. By 2014, plans were underway to develop a course for healthcare assistants, whose skills were reported to vary considerably. The Outreach team noted that some healthcare assistants did not understand the processes behind the MEWS and that they would sometimes override T3 prompts and did not effectively escalate out of range clinical observations to nurses. Likewise, nurses' skills were also reported to be varied across different specialties. These variations were reported to guide nurses' responses to MEWS scores and

thus effectively affect patient outcomes. Some nurses were reported to be more proactive, initiating early interventions and liaising with specialty medical teams, while others needed continuing prompting and support from the Outreach team. The Outreach team also reported instances of overreliance on T3 by ward nurses. They noted that this often resulted in nurses failing to recognise patient deterioration, thus delaying appropriate interventions to stop the patients' condition getting worse. They also reported lack of support for junior doctors from senior clinicians. They highlighted that junior doctors were often left to manage wards and had become heavily reliant on support from the Outreach team.

There is a risk of overreliance on T3. Nurses are now taught on the high dependency skills course, but skills utilisation on the wards is variable, as they are not expected to do such assessments. Outreach team now requires nurses to handover systematically to encourage a change of culture and reduce variability... [Prior to T3 implementation]

The Outreach service was very isolated and guidelines have improved awareness and nurses better informed. T3 has improved access and is a massive safety net for nurses, especially junior or inexperienced nurses. MEWS prompt initiation of interventions and seek senior review at ward level...

T3OutreachSister1

This was the only T3 evaluation where documentation was available. The Outreach team acknowledged the limitations of their evaluation. OutreachSister2 noted that most of the improvements reported were based on anecdotal evidence because they did not have the resources to carry out formal studies with specific focus. The Outreach team argued that the current T3 PDA, although much better than the previous one was small and limited what they could do. They instead preferred Ipads, arguing that they would improve their productivity

because they would have access to additional applications such as the British National Formulary for medicines, better viewing of chest x-rays, real time emails to colleagues for referrals and fast tracking escalations, easy access to observation trends and instant access to the MedICUs database among other advantages. They also acknowledged that both the MedICUs database and T3 were underutilised and resources available for the service were very limited. They were also keen to extend the service to cover the study Trust for 24 hours like what many NHS Trusts were doing, rather than the current provision between 8am and 6pm. This evaluation was focused on various aspects of the critical care outreach service, and not directly at T3.

#### **7.5.6 Root cause analyses for VTE**

T3LeadConsultant noted that his role included “frequently” performing root cause analyses to investigate adverse events and clinical incidents reported in the study Trust in collaboration with the Clinical Governance Department. The process of carrying out root cause analyses were based on the National Patient Safety Agency’s pathway (see Figure 5.3). Each investigation was treated and reported as an individual case. The purposes of the root cause analyses were to identify potential or actual clinical, operational and other problems that resulted in adverse events and ensure that remedial actions were taken to stop them recurring. The root cause analyses were used for the study Trust’s internal structures, such as supporting the serious the incidents group, preparing the Trust’s legal reports and for coroner’s reports. Prior to T3 implementation, these investigations were undertaken using paper-based MEWS charts and hand written patient notes were reported to be varied and

often illegible. T3LeadConsultant noted that the process of carrying out investigations prior to T3 implementation had been cumbersome and unreliable.

Operationally, you can track where patients are... you can track the number of clocks on T3, the number of critical care reviews... we went up from 20/25 to 120 reviews per week being triggered... rescued needless deaths, improved clinician confidence, improved outcomes, worked with the NHS Modernisation Agency which showed stepwise changes... I frequently do RCAs for internal reviews serious incidents group and the coroners' and legal reports... ability to actually look for and find notes, I can get from these systems without worrying where they were... often notes were handwritten, illegible, but we can now get these... charting processes, timeliness of observations, we also know that if you don't have a significant number of observations at night, nursing staff aren't actually looking at patients...

T3LeadConsultant

T3LeadConsultant argued that T3 had vastly improved access to and timeliness of clinical observations. Findings from root cause analyses were reported to have contributed to improvements in acute care service provision across the study Trust. For example, T3LeadConsultant noted that following T3 implementation, it took some time to find out that nurses were not doing observations correctly at night. NICE guideline 50 recommended that hospitalised patients should have at least two sets of clinical observations taken per day. Additionally, observations should be individualised depending on the patients' MEWS scores, clinical presentation and current treatment regimens. Root cause analyses revealed that some ward nurses were not performing clinical observations at night, arguing that they did not want to disturb patients' sleep. Findings from root causes analyses led to the observation



that several patients were deteriorating at night because key observations were being missed or not acted upon. Additionally, some nurses continued to write observations on pieces of paper, then updating them on T3 afterwards. The Outreach and hospital at night teams worked with the affected ward areas to improve education about the recording of clinical observations, correct use of T3 and escalation processes and improving understanding of NICE guideline 50. T3LeadConsultant noted “stepwise changes” following these interventions. There was awareness of the root cause analysis process amongst interview participants and general consensus that T3 had led to improvements in carrying out investigations. However, these assertions were based primarily on professional judgements because no evaluations were specifically carried out to look at the improvements that resulted from T3. Additionally, there was no overall assessment or report of the root cause analyses that had been carried out apart from in-depth case by case reports of investigations.

### **7.5.7 T3 user group network**

T3 suppliers, in collaboration with the developing and study Trust set up a user group network. The purpose of the network was to bring T3 users together to share experiences and find solutions to common problems, as well as helping to evaluate and contribute towards the further development of T3. By 2014, several NHS Trusts had adopted T3 and regular meetings were held between the user group and the supplier. However, T3LeadNurse revealed that the focus of meetings had shifted to national issues affecting member Trusts and how T3 could be used to mitigate some of the issues rather than the need of the individual Trusts in relation to T3.

It is not the usual user group, more of a help group or informal network... This is very much the software supplier presenting, ‘this is our next great thing’... very little time

to explore some of the issues [related to T3], probably a very good reason why they do that. We know where it's been implemented, being the flagship organisation. We did lots of site visits for other Trusts to come here, and we communicate with them directly... It's about gaining consensus and priorities amongst the many customers. People tend to normally agree as there will be national issues that are important at that point in time... It is more like the supplier just horizon scanning... "What will sell? What do people want? What will get us new business or keep us in business?"

T3LeadNurse

The frequency of the user group meetings was dictated by T3 supplier and the prevailing national initiatives. As noted by T3LeadNurse, the user group meetings were an opportunity to involve users in the evaluation and continued development of T3. However, opportunities were missed to learn from individual member Trusts' experiences. Instead, the meetings appear to have focused on the commercial interests of T3 suppliers and identifying commercial opportunities based on prevailing national initiatives rather than identifying and responding to the needs of the individual Trusts. Furthermore, the expansion of the user group meant that unlike the early development stages, updates to the T3 software algorithm would have to be agreed within the user group. T3LeadConsultant noted that the costs of updates and customisation work had soared as the user group had grown and it was increasingly becoming difficult to make any changes to T3. Consequently, the study Trust could not afford some additional functions that had been added to T3.

### 7.5.8 Informal evaluations

All the interview participants discussed many informal evaluations that were carried out for T3, particularly observations and professional judgments in the course of their routine work. The reported evaluations were mainly focused on monitoring the progress of T3 implementation and its usage in various departments across the study Trust. T3LeadConsultant noted that immediate improvements were reported in orthopaedics and trauma departments following T3 implementation. However, he also noted that it took longer to notice improvements in general medicine wards. T3LeadConsultant noted that “step wise” changes were only realised after the introduction of a Medical Assessment Unit in the study Trust, which relieved pressure on beds in the Accident and Emergency department as well as on general wards. However, it was not clear whether the improvements were a result of T3 or the introduction of the Medical Assessment Unit.

T3LeadNurse and the Outreach team reported a culture change regarding taking and recording of patient observations since T3 implementation. This change was attributed to ward nurses’ familiarity with the clinical assessment of acutely ill patients, which resulted in improvements in escalation processes. However, WardSister1 argued that nurses’ familiarity with T3 also brought problems such as failure to act on observations that were out of range and general complacency. She noted that these trends were essential to effectively assess patients and introduce appropriate interventions. She argued that such omissions would have been picked up using paper-based MEWS charts because they clearly showed observations trends. Furthermore, healthcare support workers performed most of the clinical observations

and some were unable to identify impending deterioration and sometimes did not escalate high MEWS scores to nurses.

The problem is when you have a system in place I don't think they [Healthcare Assistants] always look at what it tells you, they just pop observations in and at the end it says inform the nurse... sometimes they don't inform us... may lead to complacency, familiarity... The paper observation chart had a place... one of the problems with T3 is that doctors have stopped looking at the trends... they just ask for last observations and I keep asking them about the trends, which is what we used to do...

T3WardSister1

Medical cover and doctors' response times were reported to be inadequate due to changes in junior doctors' working hours, especially in the evenings and weekends. In such cases, T3ImplementationNurse argued that ward nurses were "empowered" to directly contact the patients' respective consultants or the Outreach team to escalate high MEWS scores. Furthermore, because nurses were now being trained in high dependency skills by the Outreach team, they were reported to be more confident and "just run with it" (T3LeadNurse) by starting and escalating interventions. This had reportedly led to significant improvements, although they had not been formally assessed. There was consensus amongst most interviewees that T3 had improved awareness of the MEWS scores around the Trust. The uptake of T3 and improvements in attitudes around observation taking were reported to have improved across the Trust. These improvements had reportedly made the work of the Outreach team easier. OutreachSister2 argued that T3's impact was "obvious" and did not even require evaluation. These assertions were based on informal comparisons with the

period prior to T3's implementation. Furthermore, the Outreach team's conclusions further illustrated a consistent view amongst clinical managers and decision makers that T3 was a management control system.

Prior to T3, there was no way to know what's going on... documentation was poor, sometimes illegible, often observations were recorded but not acted on, but MEWS require you to act, and now with T3, instant alerts of out of range or abnormal observations are available for the Outreach team and across the whole Trust; so how do we measure that? It's obvious, it doesn't even need to be measured, it makes our job more streamlined, less time consuming, sits very well with what we do, a picture of the whole Trust . . . real time access...

T3OutreachSister2

Some Ward based nurses and line managers raised concerns regarding their lack of involvement in the adoption of technologies such as T3. However, the Outreach team argued that the nursing hierarchy were not actively involved with T3. The Lead Outreach Nurse arranged regular meetings with nursing line managers and modern matrons to improve their involvement and address the concerns they faced on their wards. They also sought to help nurse managers to utilise the Performance Application to identify areas of improvements on their wards. Some nurses were reported to distrust T3 and did not value it as a decision support tool. Some line managers reported that T3 did not adequately cover patients in specialist areas. Some nurses and line managers argued that the patient monitoring systems they had in place in their departments were more sophisticated than T3. Some T3 fields were reported to be too rigid and instead of helping nurses, they were said to be taking away nurses' independence to make professional decisions. However, T3 was thought to be

appropriate for most general wards, which were generally staffed by inexperienced nurses without adequate skill mix.

You need it [T3] for some people, not for everyone. If you design a system to help inexperienced nurses, you can't take anything away from it as it has to be fool proof for them and the rest of us have to continue using this system but we aren't inexperienced here are we? T3 doesn't make a difference to us helping patients [in this unit]... if a patient was unwell, then we would be regularly checking their observations anyway and T3 may actually become a hindrance... often if you call a doctor to review a patient, they prioritise a high MEWS score but often unwell patients may not necessarily have a high score and you know from experience that observations alone do not reflect that... I would look more on the patient and our own monitors here. The first thing you do is [patient clinical] observations so T3 is irrelevant really, it's just a device you enter data into, but for some people it guides them...

T3WardSister1

WardSister1 also questioned the decision to adopt T3 and its value in supporting clinical processes. Instead, she felt that it was only useful for the Trust's legal obligations by ensuring that observations were kept on computerised records.

Decisions to adopt technology are probably made by senior nurses who have not been in a clinical environment for a long time... it [T3] just came and nobody asked us for opinions... It may be useful to show relatives, to prove that observations have been done, legally this can't be forged as it has both date and time stamp...

T3WardSister1

## **7.6 Discussion**

In this section, the T3 evaluations that were carried out are discussed in the context of the key aspects of evaluation that were identified in the literature review (CDSS evaluation framework). This was done to establish whether these evaluations provided enough information to the various actors involved with T3, especially the key decision makers at the study Trust.

### **7.6.1 Purposes of T3 evaluation**

The primary purposes of the formal evaluations were to show that the study Trust was compliant with national guidelines and to justify its investment in T3. However, the audits performed for Dr Foster and IQNARC were not focused on T3, but instead on the production of information that was required by the Department of Health and other regulatory bodies. These audits were also important for the Trust because they were used to rank NHS Trusts across the UK, which in turn would have an effect on the image of the Trust to various stakeholders, public perception and potentially affect funding from commissioners. The purposes of these audits were determined at national level and the Trust Board had a responsibility to ensure that the study Trust complied with the requirements. Investing in T3 was one way of showing that the study Trust was taking the necessary steps to adhere to guidelines and consequently improve patients' outcomes.

The benefits realisation studies only looked at whether the study Trust had achieved the expected benefits as defined by the developing Trust. However, it would be expected that the contexts of the two NHS Trusts were different and the benefits that were realised by the developing Trust would not necessarily be achieved by the study Trust. Additionally, T3

evaluations at the developing Trust were based on its implementation on a single ward unlike the study Trust where it was implemented across the hospital. As such, important issues that may have been relevant to the study Trust may thus have been missed. Also the second benefits realisation study focused on whether the initial benefits had been sustained seven years following T3 implementation. This information was required to support decision makers to decide whether to continue using T3 or replace it with another track and trigger system. This appears to be a limited focus for the evaluation because it did not consider other issues that may have arisen within the study Trust over the seven year period. Additionally, the outcome measures of the benefits realisation studies were based on parameters that were directly related to the critical care department but generalised across the study Trust. However, focusing on the priorities of one department over many others in the Trust may not give a balanced picture of the effects of T3 in such a big NHS Trust.

Evaluations that were carried out by the developing Trust focused on assessing the robustness of the T3 software algorithm and patient safety levels using commonly used clinical methods. These evaluations were considered important by the study Trust and partly influenced their adoption decision. However, the developing Trust did not evaluate the effects of T3 on the whole patient care system. This level of evaluation could have looked at the wider context of the environment where acutely ill patients were managed. It would ideally have involved T3 users, the differences between the clinical specialties and T3's suitability to those patient groups and its interoperability with the existing workflow and legacy electronic systems. By not evaluating these aspects, the study Trust missed an opportunity to understand the wider effects of T3 beyond the critical care department and reporting requirements of the relevant national audits. It would appear that the evaluators were primarily concerned with showing



that they were carrying out the appropriate evaluations as expected by the key stakeholders such as the Department of Health, taxpayers and commissioners. The evaluation purposes could have been widened to look at T3 effects across the study Trust rather than just the critical care aspects and those dictated at national level. However, there appears to have been a lack of motivation to commission such evaluations. The lack of motivation may have been a result of the absence of incentives to carry out evaluations beyond the minimum requirements. The lack of incentives could have influenced the study Trust's decision makers to only fund the evaluations which would produce the results that they could use, such as the Dr Foster and ICNARC audits. The limitations in the purposes of evaluations may also be because the key decision makers, such as T3LeadConsultant did not believe that any other evaluations apart from the ones that they carried out would have produced useful information.

The purposes of formal evaluations could have been widened to look at T3 effects across the study Trust rather than just the critical care aspects. Also, many claims that were made about the benefits of T3 were not backed by evidence. Some benefits were related to general service evaluation and other ongoing initiatives and not specifically T3. It was therefore difficult to attribute some of the improvements to T3.

### **7.6.2 Approaches and methods used to evaluate T3**

The scope of the evaluations that were carried out by the study Trust was limited. Like many examples identified in the literature review, these evaluations were based on positivist approaches that looked at the achievement of measurable outcomes. However, these evaluations were informal and primarily based on professional judgments at different stages

of T3's adoption, particularly in the course of implementation and routine clinical working. Apart from the benefits realisation studies, all internal evaluations were formative and based on professional judgements. Most of the benefits that were reported were primarily based on audits with a particular focus on measuring and reporting improvements to key stakeholders. However, the study Trust had no control over the methodologies of these audits. Internal studies used uncontrolled before and after designs which made it difficult to attribute causation to T3 and with little control over confounding factors that may have potentially affected the observed outcomes. Most evaluations undertaken in the study Trust were related to the provision of critical care services, track and trigger functions and patient safety aspects of T3. There was no summative evaluation at the end of the pilot or Trust wide implementation.

Evaluations such as the Dr Foster and IQNARC audits did not consider the human and organisational issues that were already present when T3 was introduced. For example, the assumption by the implementation team that T3 would work across different clinical departments ignored the differences between the various intended user groups and patients. As such, key contextual issues such as the differences between clinical workflow across the study Trust and the interrelationships between the key stakeholders were not evaluated. Although positivist approaches are widely credited for their ability to control confounding factors, the audits that were carried out were focused on wider services such as the critical care department and Trust-wide measures such as deaths, which were not directly linked to T3. It was likely that some of the improvements that were attributed to T3 could have resulted from other concurrent change efforts. Although T3 was reported to have brought "overwhelming improvements" from both benefits realisation studies, it was surprising that

three was no documentation at all available apart from superficial references in the study Trust's annual reports and quality account reports. Additionally, using the same methodology to repeat the benefits realisation study seven years following implementation either assumed that there had not been any important changes to take note of during this period or ignored them completely. The benefits realisation model emphasises the importance of identifying stakeholders and their respective responsibilities to ensure delivery of the intended benefits. It appears, however, that T3 implementation, benefits prioritisation and evaluation became a critical care department rather than Trust-wide project. The lack of involvement of other key stakeholders such as nurses who were the intended T3 users may have created barriers. Indeed some of the nurses and nurse leaders noted that T3 was imposed on them and that issues they considered important in their clinical areas had not been considered.

The majority of the formative evaluations that were carried out during the pilot and Trust-wide implementation phases were primarily based on professional judgements by the implementation team. Most of the claims made about the "overwhelming improvements" and other benefits following T3 implementation were based on anecdotal evidence. There were assumptions by the Outreach team members and T3 implementation team that T3 introduction had automatically led to improvements because the previous paper-based system had many shortcomings. Also despite the lack of evaluation documentation, the majority of the interview participants believed that the necessary evaluations had been performed appropriately by the T3 project team. This was primarily because they trusted their colleagues in the critical care department, Outreach team and the project team. As such they did not question the methodologies that were employed and the results that were reported. However, the evaluation methods that were used, such as the audits could not prove that the

reported benefits were directly attributable to T3. Instead, they focused on general service evaluation and other ongoing initiatives that were not specifically related to T3. For example, the IQNARC and Dr Foster audits were already being used before T3 implementation and the study Trust had been meeting the related national targets. Also in 2009, the Trust had undertaken remedial action with support from Dr Foster and the Strategic Health Authority to reduce death rates as part of the HSMR reporting. Similarly, root cause analyses were already being undertaken as part of the wider clinical governance framework through the existing departmental structures before T3 implementation. Although the processes of performing root cause analyses may have been improved by T3 introduction, it is difficult to attribute other claimed improvements solely to T3.

### **7.6.3 Contextual and organisational issues**

All the evaluations that were carried out in the study Trust were performed by members of the critical care department (including the Outreach team). This may explain why most of the evaluations were primarily focused on critical care services and not the study Trust in general. However, T3 was implemented in different departments across the study Trust, which had diverse groups of patients and different healthcare professionals who had varied skill sets, which may have a bearing on how it was used and perceived. For example, WardSister1 argued that MEWS scores that required urgent remedial action in a typical medical ward may be “normal” in specialist wards. She argued that arbitrary use of the MEWS across the Trust was inflexible and resulted in increased workloads in already busy wards. There appears to have been an assumption by the T3 implementation team that standardisation of processes through T3 would automatically result in improvements in patient safety and outcomes. However, none of the evaluations appeared to have considered the differences between the

various clinical departments. The formal evaluations also did not look at how the introduction of T3 would affect existing work patterns, professional responsibilities and its interoperability with existing electronic systems. For example, although alerts and prompts were designed to be vital aspects of T3, they were reported to be too frequent and in some cases resulted in delays in commencing or continuing with the necessary interventions. The discontinuation of the alert system by the Outreach team and the reported complacency by nurses and healthcare workers regarding T3 alerts and prompts revealed that the original design premise may not be practical in real clinical practice and that T3 was not used as intended. Although failures of paper-based MEWS charts were well documented, it was unclear whether the study Trust or by extension patients had benefited from T3 introduction. Evidence of T3's benefits from the developing Trust (prior to adoption) was limited to a small ward which was easier to control compared with Trust-wide implementation at the study Trust.

None of the evaluations looked at how the introduction of T3 would affect the existing work patterns, changes in responsibilities and its interoperability with existing electronic systems. The study Trust's focus on standardising clinical workflow using T3 may also have affected clinical areas that may have been already achieving acceptable patient outcomes. This appeared to have been particularly so for specialist areas that were already using early warning systems embedded into electronic or paper systems and effectively serving local needs. However, no comparisons were made between T3 and other legacy systems that were already being used in the study Trust.

### 7.6.4 Barriers to evaluation

There were a range of methodological, human, organisational, professional and technological barriers that mitigated all the key aspects of evaluation that were identified in the literature review (CDSS evaluation framework). The audits that were used as the key method of evaluation had explicit blind spots emanating from their focus on measures that were considered important to the decision makers at the time. These blind spots primarily related to non-technological issues, such as T3's effect on the intended users, lack of consideration of existing clinical workflow and the prevailing wider organisation issues. There was also no effort to address or even acknowledge these blind spots by the evaluators and the key decision makers in the study Trust. These assumptions may also have had an impact on the commissioning of evaluations, particularly the key aspects such as motivations for evaluations, methods used, who and what was evaluated and the intervals of evaluations. T3 project leaders argued that T3 had achieved its purposes as a safe and efficient track and trigger system that improved patient outcomes and also produced auditable data for the study Trust. However, apart from the audits, these claims had not been measured. It could be argued that the study Trust could not justify investing in further evaluations when they were already achieving the intended benefits, however faulty the intended benefits may have been. Furthermore, some of the key stakeholders, such as the Outreach team did not even see any reason for carrying out any evaluations because "things were much better" just by introducing T3 compared with the previous state of affairs.

T3 is very much part of what we do... before, we relied on referrals, now we have instantaneous observations... we used to write in our sheets, piles of papers, which we manually added to the database... now most things are there and all we have to do is

sync... T3 is invaluable as the existing systems, if we see patients, we tend to also do observations, and we sync into the national databases and submit straight away but sometimes we have to write and input later... there was no way to know what's going on, documentation was poor, illegible, sometimes observations were recorded but not acted on, but MEWS improved as people were required to act... now with T3, we get instant alerts for out of range or abnormal results... as the Outreach team, we get instant access, we can see the whole Trust... so how do we measure [evaluation]?, it's obvious, it doesn't even need to be measured...

T3OutreachSister1

Another barrier to evaluation was that T3 was introduced as an agent to change the habits of nurses around observation taking and recording and to improve clinical workflow. However, Garg et al. (1995) found that CDSSs that were introduced as change agents had high failure rates. The initial T3 device had many technical shortcomings that hindered its use such as the poor battery, poor layout of trends, loss of connectivity and poorly fitting stylus. Nurses and doctors immediately disliked T3 and continued to use paper charts alongside it thus resulting in increased workload and defeating the purposes of T3 altogether. Although the technical issues and unintended uses were identified during the extended pilot, Trustwide implementation went ahead regardless. Trustwide implementation could have been delayed until the necessary improvements had been effected. This could have resulted in improved "buy in" from the nurses and doctors on the wards.

Many of the key stakeholders believed that various benefits had been achieved just by implementing T3. Some did not even see the need to undertake any evaluation at all. Also many benefits that were attributed to T3 may have been caused by other ongoing initiatives

which were not taken into consideration by any of the evaluations that were carried out. Furthermore, none of the evaluations were clearly focused on T3. Instead, adherence to routine audits such as IQNARC, Dr Foster and clinical governance appears to have been attributed to T3. The evaluations that were performed only involved T3 project team members and members of the critical care services such as the Outreach team. These evaluations could have benefited from involving the wider multi-disciplinary teams across the study Trust. The benefits realisation studies only looked at whether the study Trust had achieved the expected benefits, as defined by T3 developers and evaluations undertaken in collaboration with the developing Trust. However, the contexts of the two NHS Trusts were likely to be different and this may have affected the focus and methods of evaluation. Important issues that may have been relevant to the study Trust may thus have been missed. Also the second benefits study focused on whether initial benefits had been sustained seven years following implementation to support the decision makers to decide whether to continue using T3 in the study Trust. This appears to be a limited focus because it did not consider key organisational and other changes over this period of time.

### **7.6.5 Benefits of T3 evaluations**

The main benefits of T3 evaluations were to produce information that would satisfy regulatory requirements and key stakeholders such as the government, commissioners and taxpayers. Other benefits included sharing knowledge with other NHS Trusts who were part of the T3 user group and those who looked up to the study Trust as the T3 flagship Trust. However, the lack of structured evaluations directly focusing on T3 did not help to fully understand T3 as a novel technology and its wider impact on the patient care system. Claims of T3 benefits such as minimisation of risk to patients, improved user performance, efficiency



and improved patient outcomes were not specifically evaluated. Additionally, benefits such as appropriately tracking and triggering patients based on the MEWS scores relied on the appropriate use of T3. However, the multitude of barriers resulted in various unintended uses which could have resulted in patients being put at more risk compared with the previous paper-based MEWS charts. No evaluations were carried out to identify and address the unanticipated, unwanted and potentially harmful effects of T3. Additionally, T3 was not optimally used and may not have been a justifiable use of resources as claimed by the project leaders but this was not addressed by any of the evaluations.

### **7.6.6 Overview of T3 evaluations and missed opportunities**

The T3 evaluations did not address many important issues that were identified in the literature review. The formal evaluations were based on expected benefits based on outcomes that were considered important by the developing Trust rather than those that were relevant to the study Trust context. As a “flagship Trust” where T3 was heavily customised beyond recognition from the original product, many lessons could have been learnt and shared with other potential users. However this opportunity was missed because the evaluations focused on the original benefits that had reportedly been achieved by the developing Trust. There were also assumptions that by implementing T3, most of these benefits would be achieved by the study Trust. These assumptions were based on T3LeadConsultant’s knowledge and trust of evaluations performed by his peers at the developing Trust. Evaluations could have looked at the wider effects of introducing technology into human and social systems. This could have unearthed the effects of T3 on different groups of healthcare professionals, different clinical specialties and different patient

groups. Additionally, the experiences and perceptions of the various actors involved would have added to the knowledge of effects of T3 in real clinical settings.

It would be expected that T3 brought some changes to the ways of working and ultimately patient outcomes. However, there was a perception by the key actors that T3 had largely brought positive results to the study Trust. None of the evaluations looked at potential and actual harmful, unwanted and unintended effects of T3. It was difficult to fully evaluate the evaluations that had been carried out due to lack of project and evaluation documentation that would have corroborated verbal accounts of the key actors involved in the adoption and evaluation of T3. Requests made by the researcher for this information was declined on the basis that there was an ongoing tender process to replace or retain T3 and T3LeadConsultant was looking to publish some research papers with collaborators from the developing Trust and the T3 developers. Table 7.3 shows a range of issues that were missed by T3 evaluations.

Key evaluation factors based on the CDSS evaluation framework	Evaluations carried out	Issues missed by evaluations
Purposes of evaluation should cover a range of issues, rather than focus on single issues	Main focus on measurable and auditable issues and reporting to regulatory authorities	Human and organisation factors that contribute to the identified key issues
Various approaches and methods of evaluation should be applied throughout the CDSS's lifecycle	Mainly based on formative informal usability and related technical evaluations, audits and service evaluation	Interpretive and critical studies that look at interrelationships between key stakeholders, power relations, T3 integration with workflow and existing legacy systems
Contextual and organisational issues should be considered in any evaluation activity	Focus of evaluations was primarily on the critical; care department and assumptions made that T3 effect would deliver benefits across the Trust	Evaluation did not address T3's effects in the various specialties and clinical workflows across the Trust
Benefits of evaluation should be made clear to all stakeholders	Stated benefits and those that were evaluated were different	Identify a range of benefits from evaluations that would help to improve "buy in" from intended users and key stakeholders
Barriers to evaluation should be mitigated by widening the purposes, approaches and	Most barriers were identified in informal evaluations but not addressed. There was a	Identify and address methodological, technological, human and organisational barriers. Address a range of

methods, taking into consideration	perception by the implementation team that	purposes using an eclectic range of approaches and set
contextual and organisational issues as well	resistance from intended users was an expected	clear benefits to mitigate barriers to evaluation
as highlighting the benefits of evaluation	part of the adoption process	

Table 7.3 Issues missed by T3 evaluations

# Chapter 8 Cross Case Analysis

## 8.1 Introduction

This chapter uses some of the key findings of the literature review to compare the three cases th

at were presented in Chapters 5, 6 and 7. Section 8.2 discusses the evaluations that were carried out using the CDSS evaluation framework as a guideline. Section 8.3 looks at the key contextual factors that affected the evaluations, including existing technological infrastructure, organisational readiness for CDSS evaluation, the attitudes of key stakeholders towards evaluation of computerised health information systems and the external CDSS market. Section 8.4 looks at the purposes of the evaluations and Section 8.5 discusses the evaluation methodologies that were employed. The barriers to CDSS evaluations are discussed in Section 8.6 and the benefits of evaluations highlighted in Section 8.7. In Section 8.8, the factors of evaluation that were identified in the three case studies are discussed. The limitations of the evaluations that were carried out and an overview of the chapter is given in Sections 8.9 and 8.10.

## 8.2 CDSS evaluations that were carried out

The evaluations that were carried out can be broadly classified as pre-implementation, during implementation or post-implementation. Table 8.1 shows the split between these categories for each of the three cases. As Table 8.1 shows, T1's pre-implementation evaluations included peer reviews of the VTE operational policies and validation of the paper and

software algorithms. Additionally, the T1 project team also carried out efficacy and user acceptance testing with real clinical users towards the end of its development phase. T1 evaluations that were carried out during implementation included monthly CQUIN audits and informal evaluations by the project team. T1 evaluations that were carried out following its implementation include the following audits; CQUIN, Enoxaparin and Doppler Scan effectiveness, VTE management system and the Link Nurse-led NICE Quality Standards audit. Additionally, the Haematology Department also carried out Root Cause Analyses, while various informal evaluations and professional judgments were also undertaken at various stages of its adoption.

T2 pre-implementation evaluations included validation of its clinical rules and clinical evaluations that were performed by the developers. T2LeadConsultant carried out informal evaluations through peer groups and discussions with the developers prior to its adoption at the study Trust. During T2's implementation the key stakeholders carried out informal observations and professional judgments relating to its benefits. T2 post implementation evaluations included the post market surveillance audits by the developers, a patient satisfaction survey by a local medical school linked with the study Trust and various observations and professional judgments by key stakeholders. The study Trust did not commission any formal evaluations throughout the cycle of T2's adoption. Table 8.1 shows the evaluations that were carried out before, during and post implementation of T3 in chronological order.

CDSS	Pre-implementation	During implementation	Post implementation
T1	<ul style="list-style-type: none"> <li>• Peer review of study Trust VTE operational policies</li> <li>• Peer review of paper-based VTE algorithms</li> <li>• Validation of T1 software rules and algorithms</li> <li>• T1 usability and user acceptance testing</li> <li>• Informal evaluations and professional judgments before implementation</li> </ul>	<ul style="list-style-type: none"> <li>• Monthly CQUIN compliance audits</li> <li>• Informal evaluations and professional judgments during the Trust wide implementation</li> </ul>	<ul style="list-style-type: none"> <li>• Monthly CQUIN compliance audits</li> <li>• Enoxaparin and Doppler scanning cost-effectiveness audit</li> <li>• External audit of the study Trust's VTE management system</li> <li>• Root cause analyses of hospital acquired VTE</li> <li>• Enoxaparin prophylaxis and treatment audit</li> <li>• Informal evaluations and professional judgments following Trust wide implementation</li> </ul>
T2	<ul style="list-style-type: none"> <li>• Validation of T2 software rules by the developing Trust</li> <li>• T2 clinical evaluation by the developing Trust</li> <li>• Pre-adoption informal evaluation of T2 by T2LeadConsultant</li> <li>• Informal evaluations and professional judgments before implementation</li> </ul>	<ul style="list-style-type: none"> <li>• Informal evaluations and professional judgments during implementation</li> </ul>	<ul style="list-style-type: none"> <li>• Post market surveillance by supplier</li> <li>• Patient satisfaction survey by local medical school</li> <li>• Informal evaluations and professional judgments following implementation</li> </ul>

T3	<ul style="list-style-type: none"> <li>• Technical and clinical efficacy evaluations carried out by the developing Trust</li> <li>• Informal evaluations and professional judgments before implementation</li> </ul>	<ul style="list-style-type: none"> <li>• Informal evaluations and professional judgments during pilot implementation</li> <li>• Informal observations and professional judgments during Trust wide implementation</li> <li>• IQNARC audits</li> <li>• Root cause analyses</li> </ul>	<ul style="list-style-type: none"> <li>• First and second benefits realisation studies</li> <li>• Informal evaluations and professional judgments following Trust wide implementation</li> <li>• Critical Care Outreach Team service evaluation</li> <li>• IQNARC audits</li> <li>• Dr Foster audits</li> <li>• Root cause analyses</li> </ul>
----	--	--	--

Table 8.1 CDSS evaluations carried out before, during and after implementation in chronological order



The NHS Trust that developed T3 carried out feasibility, technical, clinical and economic evaluations during its adoption cycle at the study Trust. In part, these evaluations were an important factor in making the adoption decision for the study Trust. Also, T3LeadConsultant had previously worked with the lead clinicians from the developing Trust and therefore had faith in their evaluations. Additionally, T3LeadConsultant also carried out informal evaluations in the form of discussions with the developers and other peers in collegiate groups and related national forums about its benefits. During the T3 pilot phase, the implementation team's focus was on training users and customising T3 to the study Trust's clinical workflow and existing electronic systems. The majority of the reported evaluations during this period were informal and often based on anecdotal evidence. For example, the multi-disciplinary T3 project team met every afternoon for "wash up" meetings whereby issues arising from the pilot ward were discussed. These meetings were considered important by the T3 project team because they reportedly unearthed contextual issues that would then be fed back to the developers and rectified where necessary. The "wash up" meetings continued throughout T3's Trust wide rollout phase. The T3 implementation team also reported that they carried out needs assessments of wards before T3 implementation and meetings with specialist clinical areas to establish their additional needs. They considered the needs assessments as additional evaluations of T3. It could be argued that their genuine understanding of what they considered to be evaluations could have resulted in misleading other stakeholders about evaluations that had been carried out. The following evaluations were carried out following T3's implementation at the study Trust:

- first and second benefits realisation studies
- informal evaluations, observations and professional judgements by the key stakeholders
- the Critical Care Outreach Team service evaluation
- Dr Foster audits and Good Hospital Guide mortality rankings
- IQNARC audits
- Root Cause Analyses

One year following Trust wide implementation, the T3 project team performed a benefits study, which looked at whether the study Trust had achieved the expected benefits from T3 implementation. The benefits study concluded that all the expected clinical, operational and financial benefits had been achieved. The benefits study was repeated seven years following T3 implementation in part to establish whether the originally achieved benefits had been sustained and also to inform the re-tender process for T3 or alternative replacement. The Critical Care Outreach Team also carried out a service evaluation, which included an informal assessment of T3's impact on the outreach service. Other evaluations noted by the T3 project team included the Dr Foster and IQNARC audits and the Root Cause Analysis that were undertaken in collaboration with the Clinical Governance and other relevant departments. However, IQNARC audits and Root Cause Analyses were already being carried out in the study Trust and did not directly evaluate T3. Similarly, the Dr Foster audits were focused on analysing and reporting the number of deaths while patients were hospitalised and were not directly focused on T3. None of the three CDSSs had a systematic evaluation portfolio. In some cases, key

stakeholders were unaware of the evaluations that had been carried out by the project teams and vice versa. This was particularly so for T1 and T3, where users were unaware of the ranges of evaluations that had been performed. This was unexpected for T1 because departmental monthly VTE audits were published on the study Trust's intranet and discussed in multi-disciplinary team meetings on a regular basis.

This section has shown that the range of evaluations that were carried out for the individual CDSSs were formative and mostly focused on their successful implementation. The evaluation portfolios did not consider the different types of evaluations that could be carried out at different phases of the CDSSs' lifecycles. Additionally, the lack of linkages between the various evaluations that were carried out for the individual CDSSs and across cases did not allow for transfer of information to inform the key stakeholders and the evaluations themselves. Such transfer of information would have enabled learning from prior evaluations and contributed to the usage and ultimately decision making regarding the CDSSs.

### **8.3 Key contextual factors affecting evaluations**

All of the case material was gathered in the same hospital so at least in one respect it is true to say that the evaluations were conducted in the same context. However, some important contextual differences were identified between the individual clinical specialties involved. These differences include:

- variations in existing computersied health information system infrastructures across the study Trust
- wider NHS and healthcare technological contexts in the form of national guidelines, professional bodies and peer groups
- existing culture of CDSS evaluations within the study Trust
- attitudes of the various members of the multidisciplinary teams towards evaluation of CDSSs

Each of these differences will now be considered in turn.

### **8.3.1 Assessment of existing technology infrastructure**

An assessment of existing legacy systems infrastructure was identified in the literature review as being critical to successful evaluation of new CDSSs. This assessment would help to identify and optimise existing resources across the NHS and regulatory authorities, as well as at organisational and individual levels. Table 8.2 shows the factors that need to be considered when assessing existing CDSS infrastructure to support evaluations.

Key dimensions	Sector wide factors	Organisational factors	Individual factors
<ul style="list-style-type: none"> <li>Attitudes of the key stakeholders</li> </ul>	<ul style="list-style-type: none"> <li>identifying the enablers or inhibitors of evaluations in the sector and related networks</li> <li>assessing the wider NHS visions and initiatives</li> <li>assessing the existing political economy of evaluations</li> </ul>	<ul style="list-style-type: none"> <li>assessing the organisational culture and motivation for evaluations</li> <li>identifying the influences from peer networks</li> </ul>	<ul style="list-style-type: none"> <li>identifying the key stakeholders</li> <li>Understanding the characteristics and attitudes of key stakeholders towards evaluations</li> </ul>
<ul style="list-style-type: none"> <li>Conditions necessary for evaluations</li> </ul>	<ul style="list-style-type: none"> <li>identifying existing and making use of sector-wide regulatory, structural, systemic requirements that support evaluations</li> </ul>	<ul style="list-style-type: none"> <li>identifying and making use of the existing organisational infrastructure to support evaluations</li> </ul>	<ul style="list-style-type: none"> <li>assessing the capabilities of the various stakeholders within the organisation and how they could influence evaluations</li> </ul>
<ul style="list-style-type: none"> <li>Resources required to facilitate evaluations</li> </ul>	<ul style="list-style-type: none"> <li>identifying and mobilising existing external resources that may support evaluations</li> </ul>	<ul style="list-style-type: none"> <li>assessing the organisation's readiness to undertake the evaluations, e.g. human, structural, financial and non-tangible resources</li> </ul>	<ul style="list-style-type: none"> <li>identifying and utilising individuals within the organisation who have the required knowledge and skills to design and carry out evaluations</li> </ul>

Table 8.2 Factors to consider when assessing existing technological infrastructure to support CDSS evaluations

Variations were noted in the existing legacy systems' infrastructures across the different departments where the three CDSSs were implemented. Although there were a number of legacy systems that were being used across the study Trust prior to the adoption of the three CDSSs, none of these systems were specifically designed to support VTE risk assessment, management of prostate cancer patients or the assessment of acutely ill patients. T1 and T3 were implemented to directly replace paper-based risk assessment tools. They also introduced new ways of working and new responsibilities to some stakeholders, while also taking away responsibilities from other stakeholders. The existing legacy systems in these areas were not being used to directly support the services for which T1 and T3 were eventually developed to support. In contrast, the Uro-oncology specialist nurses were already routinely using existing computerised legacy systems supported by paper-based clinical pathways prior to T2 adoption. Although no formal evaluations had been carried out with respect to these legacy systems, the urology team were familiar with the electronic processes and were in a position to make comparisons between the previous systems and T2 as well as informal evaluations and professional judgments. In contrast, T1 and T3 users were using electronic technologies for these processes for the first time so the degree of CDSS novelty would have been greater than that of T2 users. However, there was no evidence of formal evaluations of any of the existing legacy systems on which the respective CDSS project teams could learn from.

The decisions to adopt all three CDSSs were made on the assumption that they would be integrated seamlessly with the legacy systems that were already in use in the study Trust. However, by the end of 2014, full integration had not occurred. For T2, users were using two or even three disparate systems, while some T3 users were now concurrently using the "old" paper-based charts to complete routine tasks. T1 and T3 were integrated with the existing

clinical results reporting system to varying degrees. All three project teams sought to assimilate the respective CDSSs into routine clinical practice, thereby redesigning clinical processes in accordance with the NICE guidelines. This was achieved to varying extends by the T1 and T3 project teams. They sought to integrate the respective CDSSs with existing electronic legacy systems in order to try to encourage their use while also monitoring compliance using existing administrative parameters that were embedded within the legacy systems. Only the T1 project team formally undertook a systematic assessment of existing legacy systems infrastructure prior to adoption. This resulted in T1 being developed and integrated onto the study Trust's clinical results reporting system platform, which was widely used by clinicians, who in this case were the targeted users and ultimate clinical decision makers. This platform also enabled the T1 project team to carry out formal evaluations using existing parameters in the clinical results reporting system and to create management reports. Additionally, the T1 implementation team were able to identify and deal with specialty specific factors that would have been difficult to identify using the paper-based system. In contrast, T2 and T3 encountered various integration problems with the clinical results reporting system. In turn, the lack of integration with legacy systems affected how they were both used in practice and also limited the opportunities for evaluation. However, these issues were not addressed by any of the evaluations that were carried out.

This sub-section has shown that:

- an assessment of existing legacy systems infrastructure is essential for CDSS evaluation
- an assessment of non-computerised decision support systems helps to understand the wider effects of CDSSs

- previous evaluations that have been carried out with respect to existing legacy systems and non-computerised decision support systems should be drawn on to inform CDSS evaluations
- the potential effects of CDSSs on existing workflow and users responsibilities should be assessed
- CDSS interaction with existing legacy systems, non-computerised systems and clinical workflow should be assessed

### **8.3.2 External pressures for CDSS evaluations**

Since the early 2000s, there had been a drive by regulatory authorities, particularly NICE, with support from professional groups for NHS Trusts to adopt CDSSs to support clinical practice and improve clinical decision-making. All three CDSSs were based on NICE guidelines. It could be argued that the pressure exerted on NHS organisations to fulfil these guidelines led to the adoption of these CDSSs. However, there were no clear recommendations regarding how to evaluate the technologies. Also because of the novelty of these systems and their limited commercial availability on the UK market, there was little evaluation information to inform the project teams.

The VTE CQUIN framework required monthly reporting on the number of valid VTE risk assessments that were carried out. T1 enabled the VTE project team to collect and report VTE CQUIN data electronically as well as identifying and addressing problems in poorly performing clinical areas. T3 project leaders were also required to report the study Trust's performance with respect to numbers of deaths and Accident and Emergency waiting times amongst others. Background data collected from T3 were used to show improvements in these areas as well as



supporting ICNARC and Dr Foster audits. Root Cause Analyses were also carried out with respect to patients who developed VTE during or following hospitalisation. However, Root Cause Analyses had no specific correlation with T3. There were no regulatory requirements for the urology department to report their performance in relation to T2. This may have resulted in the limited range of evaluations that were performed for T2. All three CDSSs also brought greater control and reporting capabilities to support regulatory requirements. This was particularly so for T1 and T3 because they were under direct regulatory pressure to report their performance on specific aspects that were related to the respective CDSSs. External pressure played a significant part in determining the range of evaluations that were carried out for the CDSS albeit with a different focus. Evaluations such as the VTE and Dr Foster audits required system-wide evaluations rather than CDSS specific evaluations. It could be argued that external evaluation pressures can create confusion between CDSS-specific evaluations and broader evaluations of clinical management systems or services.

This sub-section has shown that there were external pressures from regulatory bodies and commissioners to adopt CDSSs. Although there were no corresponding recommendations on how to carry out CDSS evaluations, regulatory reporting requirements affected the range of evaluations that were carried out. CDSS evaluators need to be aware that externally driven evaluations may not necessarily reflect or serve the best interests of the organisation. There is need to carry out separate evaluations that look at CDSS effects on the organisation because externally driven evaluations may not be sufficient.

### **8.3.3 The attitudes of key stakeholders towards evaluations**

The attitudes of the key stakeholders towards evaluation of the three CDSSs varied widely. Although these differences were likely to affect how the CDSSs were evaluated, they were not addressed by any of the evaluations that were carried out. The following factors affected the attitudes of the key stakeholders towards CDSS evaluations:

- their characteristics, experience and backgrounds
- previous evaluations that they had been involved in
- their role with respect to the CDSSs
- perceived usefulness of the CDSSs
- the effect of the CDSSs on their way of working

Each of these factors will now be discussed in relation to each CDSS.

The attitudes of the key stakeholders involved with T1 evaluations varied across professions and clinical specialties. Although the Haematology department was in charge of the overall management of VTE across the study Trust, each department had its own approach to the implementation of the NICE VTE guidelines to suit their respective patient groups and additional guidance from their specialty professional groups. This was particularly relevant for specialist departments such as the cardiology, stroke and neurology services that widely used thromboprophylaxis treatments before T1 was implemented. Their experiences prior to T1 introduction informed their attitudes and understanding of VTE as an important clinical problem. Senior doctors and nurses in these departments hailed T1 as one of the most successful IT projects that had been implemented by the study Trust. However, despite the positive feedback, there were concerns by some key stakeholders in these departments that

the rigid nature of T1 data collection and the resultant monthly VTE CQUIN audits were not representative of the complexities in their departments. Some key stakeholders argued that VTE CQUIN audits were designed to produce management data and were not geared for looking at patient outcomes and clinical processes that supported these outcomes. The T1 project leaders reported meeting challenges in obtaining “buy in” from some specialties, especially the surgical departments. However, they noted that once clinicians realised that it was compulsory to use T1 and that penalties would be incurred if they did not do so, they subsequently changed their attitudes. Consequently, compliance was reported to have improved but some clinicians were reported to be “just doing the minimum that was required” to result in what would be considered a valid VTE risk assessment. A significant number of clinicians also deferred the VTE risk assessment process as many times as the system allowed them to without incurring penalties. In cases where they could no longer defer, they were reported to do “just enough to tick the box” (T1ImplementationNurse). These practices were not picked up by the monthly VTE CQUIN audits and were instead identified by VTE specialist nurses during routine follow ups, through monitoring of the deferrals and sometimes through Root Cause Analyses.

T2 nurses were the only user group that had previously used computerised health information systems for their routine work prior to CDSS implementation. However, no formal evaluations had been carried out despite their assertions that the legacy systems were better than T2. T2Nurse1 noted that T2 did not add any value to her clinical decision making, noting that it was now more cumbersome to run the clinics using three disparate systems, including T2. To get around these complexities, the T2 nurses reported finding “ways around” T2, whereby they would input just enough data to obtain the GP outcome letter. They were aware that as long

as they saw the required number of patients in the T2 virtual clinic, as measured by the number of GP outcome letters produced, then their managers and commissioners would be satisfied. Consequently, it would appear to decision makers that T2 was being appropriately used. However, T2Nurse2 argued that despite the additional workload and lack of T2 integration with existing electronic systems, it was an important tool to reduce the risk of harm to patients and also an essential aid to improve nurses' clinical decision making. T2Nurse2 noted that despite the fact that all the Uro-oncology specialist nurses were highly experienced, there was no guarantee that they would all follow the NICE guidelines in the same way. He argued that T2 would thus be a "safety net" to counter human errors and to ensure uniformity of decision-making. He also argued that it was too early to evaluate T2 and that it should be given time to settle down before looking at ways to evaluate the benefits to the study Trust.

In similar fashion to T1, the attitudes of the key stakeholders towards T3's evaluation varied between the various departments and healthcare professionals. Specialist departments that had traditionally managed the care of deteriorating patients saw T3 as an unnecessary addition to what they were already doing. In contrast, less experienced departments viewed it as an opportunity to upskill their staff's performance. However, the evaluations undertaken did not address the differences between the departments across the Trust. There appears to have been a general assumption by T3 project leaders that every department would benefit from the CDSS without assessing their existing capabilities and the disruptiveness that could be caused by T3 introduction. Additionally, the shift in responsibilities from clinicians to nurses and their readiness to adopt new ways of working were not explored in any of the evaluations. There was a general feeling amongst some key stakeholders that T3 had been imposed on them.

The majority of existing computerised health information systems (including CDSSs) in the study Trust were mostly used by clinicians and a small but growing number of nurses, particularly in specialist areas. Most nurses were reported to exhibit negative perceptions and scepticism about the clinical value of computerised health information systems in general, and specifically the three CDSSs. All the three project teams reported that most nurses did not routinely use these systems. In contrast, most doctors considered the study Trust's clinical results reporting system and related computerised health information systems to be useful and complimentary to their clinical roles. Additionally there was a general perception amongst doctors that the study Trust's legacy systems and CDSSs, particularly T1 and T3 were superior and better integrated than those used in comparable regional hospitals where they had previously worked. Consequently, the adoption of all three CDSSs was reported by the key stakeholders to have improved the study Trust's standing as a leading and innovative teaching centre in the region. However, there was awareness amongst most senior clinicians, managers and some CDSS users that both T1 and T3 were under-utilised. Often they cited lack of capacity, conflicting priorities and other operational barriers as the causes of such limitations. Although nurses and clinicians in the urology outpatients' department extensively used the existing legacy systems prior to T2 adoption, T2 was also reported to be under-utilised. There was uncertainty about CDSS evaluations for all three cases, even for those stakeholders who were involved in the development/adoption and evaluation of the CDSSs. These uncertainties related to questions about what needed to be evaluated, how to evaluate and who was responsible for evaluations. Many interviewees reported having to "cover their backs" or "covering the Trust's back", which often meant doing just enough to "tick the box" and be seen to be doing the right things. This also applied to senior clinicians and senior managers who

were keen to be seen to be complying with national guidelines and be seen to be accountable professionally, ethically and legally to the key stakeholders such as the government, taxpayers, commissioners and professional bodies. There was a shared belief by most of the key stakeholders that things were generally better following the implementation of the three CDSS. However, these beliefs were mainly based on personal and professional judgments and were not backed by evidence from formal evaluations. It would also appear that there were some vested interests amongst some of the key stakeholders who were keen to show that the CDSSs had been successfully implemented and that the expected benefits had been realised. Some of the key stakeholders, particularly the Lead Consultants and the ICT department were in a position to exert significant power to ensure compliance and usage of the CDSSs. However, this did not necessarily result in acceptance of the CDSSs.

This Section has highlighted that the key stakeholders' vested interests can result in less effective evaluations because they may primarily focus on protecting their interests. Also, the level of user engagement in the evaluations was variable. Indeed, in the majority of evaluations, there was minimal, or no user engagement at all. The majority of the key stakeholders argued that the primary evaluations were designed to "tick the box" and protect the interests of senior consultants and Trust executives rather than improving their work processes and patient outcomes. This raises questions about how much these evaluations reflected the real effects of CDSSs in these clinical areas.

## **8.4 Purposes of evaluations**

The purposes of the evaluations that were carried out at the different stages of the CDSSs' adoption are summarised in Table 8.3. Pre-implementation evaluations primarily focused on

assessing the clinical and technical efficacy of each CDSS. For T1, these evaluations included the validation of software algorithms, clinical evaluation and interface and user testing. T2 developers carried out similar evaluations to those that were carried out for T1 in collaboration with the developing Trust. Prior to its adoption in the study Trust, T3 underwent various evaluations at the developing Trust. These evaluations were focused on outcomes such as T3's technical and clinical efficacy, cost effectiveness, guideline adherence and reducing the rates of mortality. Table 8.3 shows the purposes of CDSS evaluations at different phases of their adoption lifecycle.

T1	T2	T3
<p>To ensure appropriate translation of NICE guidelines into practice</p> <p>To ensure appropriate translation of the Trust's VTE policy</p> <p>To assess efficacy of the tool's software</p> <p>To check usability of tool interface and rules with real clinical users</p> <p>To check compliance with VTE CQUIN targets</p> <p>To assess the cost-effectiveness of Doppler scans, prophylactic and treatment does of Enoxaparin®</p> <p>To test the robustness of the Trust's VTE management systems</p> <p>To establish causes of VTE in hospitalised patients</p> <p>To audit Trust performance against NICE quality standards for VTE</p>	<p>To test system's alert levels and safety parameters</p> <p>To evaluate system's accuracy in making guideline adherent management plans compared with humans</p> <p>To assess the merits of T2 before adoption</p> <p>Inspecting system's accuracy and checking patient safety</p> <p>To establish patient satisfaction with telephone clinic and internet access</p>	<p>To ensure full implementation of T3 and resolve issues arising</p> <p>Assess whether expected benefits had been realised</p> <p>To assess the effectiveness of the Critical Care service</p> <p>To monitor the effectiveness of acute care services and deaths rates in the Trust</p> <p>To evaluate the uptake and impact of the Outreach service</p> <p>To assess the cause of adverse events in hospital and put in place remedial actions</p> <p>To assess the effects of T3 on clinical processes and existing systems</p>

Table 8.3 Purposes of CDSS evaluations at different phases of their adoption lifecycle



Only the T1 project team carried out formal evaluations during its implementation. Informal evaluations were performed for all three CDSSs during and after the implementation phases. T1's post implementation evaluations primarily focused on maintaining CQUIN compliance and benchmarking the study Trust with other NHS Trusts. Post implementation evaluations for T2 included a 'post-market surveillance' audit that was carried out by the developers as required by the MHRA and an exploratory patient survey funded and performed by academics from a medical school linked with the study Trust. However, none of the T2 project members at the study Trust were aware of the 'post-market surveillance' audit (including T2LeadConsultant). Only T2LeadCnsultant was aware of the patient satisfaction survey. The primary purposes of formal post implementation evaluations for T3 were to assess whether initial benefits had been sustained. For all three CDSSs, the primary evaluation purposes were limited to the stated project objectives, which were to meet the NICE guidelines and to satisfy regulatory authorities. Some additional evaluations were also undertaken to assess improvements in clinical processes and patient outcomes. Examples include the root cause analyses for hospital acquired VTEs, patient safety audits and the T3 benefits realisation studies looking at patient outcomes, as well as operational and financial benefits. However, these evaluations did not clearly focus on the respective CDSSs, and instead looked at the wider clinical services that were supported by the CDSSs. The benefits realisation studies also sought to justify the technology investment by proving that the expected outcomes had been achieved and maintained. Other evaluations, such as the Critical Care Outreach Team service evaluation and the T2 patient survey sought to assess the existing services and identify ways in which they could be further improved using the respective CDSSs. However, it would appear that most of the evaluations were not consistently viewing the same system because the boundaries of the

systems kept changing in line with organisational, clinical and regulatory priorities. Additionally, there were a range of complex interrelationships between various stakeholders who were directly and indirectly involved with the CDSSs.

This section has shown that the focus of evaluations should correspond to the different phases of the CDSSs' lifecycle. Caution should be taken when carrying out CDSS evaluation as part of wider service evaluations because it becomes difficult to separate and measure the effects of different interventions. Consequently improvements or other consequences may be wrongly attributed to CDSSs without clear evidence of correlation. The complex interrelationships between CDSSs and various stakeholders as well as workflow integration should be explored. The boundaries of CDSS evaluations may also change in line with the constantly shifting organisational and regulatory priorities.

## **8.5 Approaches and methods of evaluations used**

All three CDSSs were subject to both formal and informal evaluations. These are summarised in Table 8.4. Formal evaluations were primarily in the form of clinical and technical validations of the respective CDSSs' software algorithms and underlying guidelines. Informal evaluations were often in the form of routine observations and professional judgments by the key players who were involved with the CDSSs in the course of their routine work. For example, all three Lead Consultants reported that their respective CDSSs had led to substantial improvements in practice and patient outcomes compared with the state of affairs prior to their adoption. Their adoption was often interpreted as resulting in automatic improvements in patient outcomes because of the improved accessibility of information and in some cases, improvement in data collection and reporting processes.

T1	T2	T3
Peer review of the Trust's revised VTE operational policies	Clinical validation	Informal observations and assessments during the pilot and Trust-wide implementation
Peer review of paper-based VTE algorithms	Comparison of human and computer decisions	Benefits realisation studies
Validation of software rules and validation	Assessment of presentations and developer's literature	IQNARC audits
VTE tool usability and user acceptance testing	Audits	Dr Foster audits
VTE CQUIN compliance audits	Patient survey	Outreach service evaluation
Enoxaparin® and Doppler scanning cost-effectiveness audit		Root cause analyses
External audit of the Trust's VTE management system		Informal evaluations
Root cause analysis of hospital acquired VTE		
Enoxaparin prophylaxis and treatment audit		

Table 8.4 Evaluation approaches and methods used

All three CDSSs were deemed to have been successfully adopted on the basis that they had achieved their stated objectives, even though they had not been evaluated to that effect. It also appeared that the key stakeholders did not make any distinctions between the formal and informal evaluations and those conclusions were made without the necessary evidence to back them up. It would also appear that the fitness for purpose of the evaluations that were carried out was not a key consideration as long as the desired results were achieved. For example, the methodologies of data collection and reporting for VTE CQUIN audits were changed several times to match the changing regulatory requirements without any corresponding

improvements in clinical processes or patient outcomes. Although the T1 project leaders showed awareness of this disparity, their focus at the time was to ensure that the study Trust secured the CQUIN incentives. T1LeadConsultant also noted that the T1 project team had “largely cracked the nut” by securing funding for the T1 project and producing the monthly VTE CQUIN audit reports. He noted that evaluations that would be “clinically meaningful to patients” would be carried out in the future. No evaluations were commissioned for T2 by the study Trust. In part, T2LeadConsultant was satisfied with the methodologies that had been used by the developers, arguing that there was “no need to evaluate again because I would do exactly the same (methodology) as they did”. Similarly, T3LeadConsultant was satisfied with the evaluation methodologies that had been used by the T3 developing Trust and hence only carried out evaluations to assess whether the study Trust had achieved the same benefits as those reported by the developing Trust. T2 and T3 had been co-developed by leading consultants in their respective fields and both T2LeadConsultant and T3LeadConsultant had confidence in the robustness of their peers’ evaluations. However, while these evaluations might have been adequate at the developing Trusts and also at the time of development, the context of the study Trust was different and would have benefited from new evaluations that took into consideration the local contextual factors.

Another important evaluation factor was that the novelty of these CDSSs made them challenging to evaluate. This was evident in the limited evaluation portfolios and evaluation methodologies that were applied for all three CDSSs. Additionally, for all three CDSSs, decision makers in the study Trust were primarily looking to address the immediate problems such as data collection, monitoring and reporting of compliance with national targets rather than the wider implications of the CDSSs’ adoption. However, some of the methodologies that were

applied, such as the audits and informal service evaluations were not geared to address other contributory issues that may have affected the observed outcomes. For example the monthly VTE CQUIN audits did not pick up disparities in practices that were identified by the VTE Link Nurse audit because the former was designed to just record the number of valid VTE risk assessments that had been performed within the stipulated timeframes. In contrast, the patient safety audits that were performed by the VTE Link Nurses highlighted poor usage of T1. This could be attributed to inadequate training and shortcomings in the implementation process. These issues were also not picked up by the monthly audits. In fact, the methodology that was used for the Link Nurse audit revealed that VTE compliance was about 50%, and not above 95% as reported by the monthly VTE CQUIN audits.

Apart from the VTE risk assessment audits, the majority of the evaluations were only disseminated to a few members of the project teams and senior management. Some key evaluations such as the T3 benefits studies were not disseminated outside the senior members of the project team, apart from references in the study Trust's annual quality accounts. The validity of the range of evaluations that were carried out could be open to criticism because of their informality. This was particularly so for those evaluations which were based on key stakeholders rationalising their personal and professional judgments as well as the checklists and informal audits.

This Section has shown that the evaluation methodologies that were used widely varied. For any evaluation activity, there is need to:

- consider the timing of evaluations and match with appropriate methods
- establish what system level is to be studied and match to appropriate methods

- distinguish between formal and informal evaluations
- question the level of evaluators' neutrality (whether from the project team, other parts of Trust or external to the Trust)
- consider evaluations from outside the Trust and decide whether follow-up in-house evaluations are necessary
- test assumptions regarding the perceived or "sold" benefits of CDSSs
- carry out evaluations beyond the achievement of stated objectives
- document and disseminate evaluations to all key stakeholders (including CDSS users)

## **8.6 Barriers to evaluations**

A range of human, methodological and organisational barriers were identified within the individual CDSS evaluations. These barriers are summarised in Table 8.5 and will now be discussed in turn.

CDSS	Barriers identified
<b>T1</b>	<ul style="list-style-type: none"> <li>• Lack of funding to undertake evaluations that would be “clinically meaningful to patients”</li> <li>• Assertions by some key stakeholders that T1 had been a success without the necessary evaluations to support</li> <li>• Perception by some users that T1 was bureaucratic diktat without clinical benefits for patients</li> <li>• Perception by some key stakeholders that CQUIN framework was target driven and did not have any patient benefits</li> <li>• Top down approach to both T1 project and evaluations carried out</li> <li>• Audit results not accepted by some key stakeholders as representative of real clinical environment</li> <li>• Acceptance by project leaders that evaluation methods were limited</li> <li>• T1 used as an agent to change the behaviour of users</li> <li>• T1 used to restructure clinical workflow</li> <li>• T1 fit in some key clinical areas not formally evaluated</li> </ul>
<b>T2</b>	<ul style="list-style-type: none"> <li>• Project leaders did not see the need for evaluations</li> <li>• Project leaders satisfied with evaluations carried out by the developing Trust</li> <li>• T2 registration with MHRA and CE approval reassured project leader of T2 robustness</li> <li>• Notion by key stakeholders that T2 required time to “settle” in the study Trust before undertaking evaluations</li> <li>• Lack of motivation to evaluate</li> <li>• No regulatory pressure to evaluate T2</li> <li>• Lack of evaluation customers due to reorganisation of primary care services</li> <li>• T2 used as an agent to change the behaviour of users</li> <li>• T2 used to restructure clinical workflow</li> <li>• T2 fit to the study Trust not formally evaluated</li> </ul>
<b>T3</b>	<ul style="list-style-type: none"> <li>• Lack of evaluation documentation</li> <li>• Assertion by some key stakeholders that there was no need to evaluate T3</li> <li>• Primary focus of evaluations on achievement of the “sold” benefits</li> <li>• Evaluations based on professional judgments rather than structured methods</li> <li>• T3 used as an agent to change the behaviour of users</li> <li>• T3 used to restructure clinical workflow</li> <li>• T3 fit in some key clinical areas not formally evaluated</li> </ul>

Table 8.5 Barriers to CDSS evaluations

### **8.6.1 Human barriers to evaluations**

One of the main barriers was that some key stakeholders did not see the need to evaluate the CDSS concerned at all. This was particularly so for T2 and T3 because the Lead Consultants were satisfied with the evaluations that had been performed by their peers at the respective developing Trusts. Another barrier was that the project teams made assumptions based on their professional judgment and anecdotal evidence that the CDSSs had led to improvements in the services that they supported and ultimately patient outcomes. The main basis of these assumptions was that the CDSSs made it easier for clinical leaders and managers to have access to data showing the performance and compliance of their clinical teams. For example, the audits that had been carried out for the previous paper-based VTE risk assessment tool had been difficult to complete and did not provide enough information for decision making. T1 resolved these issues by providing various compliance and performance reports at various levels such as by department or consultant of specific patient group. Likewise, T2 resolved the issues around disparate clinics that were held by the different urology teams by grouping together patients into a single database. There were also plans to utilise T2 in future to identify patients who were suitable for clinical trials. Similarly, T3 provided a platform to view the performance and compliance of clinical teams and enabled the Critical Care Outreach Team and clinicians to prioritise their work and initiate interventions timeously. Although there may have been improvements in the various aspects noted by the three project teams, no formal evaluations were carried out to confirm these improvements. Furthermore, it was difficult to ascertain whether all the positive outcomes that were attributed to T3 could be entirely separated from other ongoing initiatives across the study Trust. These assumptions were primarily based on improvements in the ability to change processes to meet regulatory



requirements but failed to identify and address operational problems faced by CDSS users whose contribution to their evaluation could have brought a unique perspective based on their experiences.

For all three CDSSs, the main evaluations were conducted by the Lead Clinicians and their respective project teams and excluded other healthcare professionals. Failing to include all key stakeholders in the evaluations may have impacted on the ownership of evaluations that were carried out. The monthly VTE CQUIN audits were produced by the study Trust's Information Department with support from the leading members of the T1 project team (T1LeadConsultant, T1LeadNurse and T1Developer). T2LeadConsultant primarily carried out informal evaluations for T2 with support from the key members of its development team. T3LeadConsultant carried out formal evaluations for T3 with support from the developers and the Critical Care Outreach Team Lead Nurse. The main evaluators for all three CDSSs were the key decision makers who had championed their adoption and had also taken leading roles in their implementation at the study Trust. The involvement of the same players at every stage of the CDSSs lifecycle raises questions around their ability to carry out objective evaluations. None of these stakeholders showed any concern around self-evaluation. Some external evaluations were carried out for T1 and T2. However, the Strategic Health Authority evaluation that was commissioned for T1 was not reported to the project team and did not affect decision-making. The T2 patient satisfaction survey was exploratory and also did not appear to have had any effect on decision-making within the study Trust. Apart from the monthly VTE CQUIN audits, the formal evaluations were not widely disseminated. Only T1LeadConsultant and T1Developer were aware of the clinical and technical validations that were carried out during the early stages of its lifecycle. Likewise, only T2LeadConsultant and the T2

development team at the developing Trust were aware of the technical and clinical evaluations undertaken during its development. T2LeadConsultant was not aware of the post market surveillance audits that had been carried out by the T2 developers. Only the key members of the T3 project team were aware of the evaluations that were carried out at the developing Trust although these evaluations had been published in professional and academic journals. Despite the lack of dissemination of the evaluations, the majority of the interviewees believed that the necessary technical and clinical evaluations had been carried out as part of the CDSSs' development.

This sub-section has shown human barriers around CDSS evaluations relating to the following:

- key stakeholders not seeing the need for evaluations
- lack of scrutiny of evaluations carried out by peers
- CDSSs primarily used as management information systems and agents for change
- CDSS project leaders' personal and professional judgements influencing and at times replacing formal evaluations
- lack of distinction between effects of CDSSs and other initiatives which were implemented simultaneously
- decisions regarding how much funding was allocated for evaluations
- the position taken by project leaders regarding evaluating their own projects
- challenges around building ownership of evaluations by not involving CDSS users
- how to effectively evaluate evaluations and ensure that they improve the use of CDSSs

## **8.6.2 Methodological barriers to CDSS evaluations**

For all three cases, there was regulatory pressure to adopt CDSSs and related systems to improve processes and clinical decision-making. However, there were no corresponding recommendations regarding how to evaluate these technologies. It therefore was up to the respective project teams to devise evaluation methods that they considered most suitable. T1 evaluations were limited to the study Trust's immediate requirements for CQUIN VTE audit data and to obtain the attached financial incentives. However, the T1 project team showed awareness that achieving the CQUIN VTE targets did not necessarily imply the correct usage of T1. The patient safety audit that was carried out by T1LeadNurse and T1WardSister also showed wide variations in compliance between the CQUIN VTE audits and the actual decisions that had been made in practice. There were no immediate evaluation requirements (from regulators) for T2 and hence no formal evaluations were commissioned by the study Trust. T3 evaluations were focused on realising benefits that had been proven by the developing Trust. By "taking the word" of their peers at the developing Trusts, the Lead Consultants for both T2 and T3 showed the significant influences that were exerted by peer groups and the effects thereof on the range of evaluation methodologies that were employed. The influences of the peer groups appear to have limited the portfolio of evaluation methodologies because clinicians who tended to use evaluation methods, which they were already familiar with, carried out the majority of the evaluations. It would appear that the clinicians at the study Trust were not critical of the evaluations that had been carried out by their peers because they also believed that these were the best methodologies to use and indeed would have also evaluated the CDSSs in the same way. Similarly, other key stakeholders in the study Trust trusted the judgment of the lead clinicians who were tasked with developing and/or adopting

these CDSSs. However, this calls into question the validity of these evaluations, especially their fit for the CDSSs in the context of the study Trust. By applying the commonly used methodologies, these evaluations had explicit blind spots because their substantive focus leaned more towards the technical and clinical validations at the early phases of the respective CDSSs' lifecycle and ignored the implementation and post implementation stages. T1 and T3 were implemented across the study Trust with little regard to the differences between disparate clinical specialties. The novelty of T1 appears to have been a barrier to its evaluation because it was the first such CDSS to be developed and implemented in the NHS. The T1 project team noted that they did not have any reference points to draw on and had no choice but to 'self-evaluate'.

The sub-section has revealed the following methodological barriers to CDSS evaluation:

- there was paucity of information regarding how to evaluate CDSSs
- evaluation methods were selected for their ability to generate information that would be reportable to regulatory authorities and to gain incentives from commissioners
- evaluation methods were not geared to assess CDSS effects on users and patient outcomes
- opinions within peer groups had significant effect on the range of evaluations that were carried out
- little effort made to learn from previous CDSS evaluations or to consult wider evaluation literature
- lack of scrutiny of peer evaluations and recommendations from collegiate and regulatory bodies

- fit of evaluation methods not assessed
- explicit blind spots of evaluations not addressed

### **8.6.3 Contextual and organisational barriers to CDSS evaluations**

The main organisational barrier was that the evaluation priorities kept changing, especially for T1 and T3. Initially, the T1 Project team excluded some patient cohorts in order to help to meet the 90% compliance to secure CQUIN funding. This was followed by blocking access to the clinical results reporting system where risk assessment was not carried out within the initial 12 hours of admission. Additional methodological changes and patient cohort exclusions were made to meet the new requirement for 95% compliance at the beginning of 2014. The T3 project team utilised T3 to meet the requirements of ICNARC and Dr Foster audits even though they were not directly related. It would appear that internal and external priorities dictated the focus of evaluations that were carried out. The top down approach to evaluations may also have restricted other evaluations outside the CDSS project teams. For example, T3 evaluations were directly intertwined with the interests of the Critical Care Department, who had led its adoption in the study Trust. T3 evaluations were therefore automatically deferred to the Critical Care Department. Unlike T1 and the T3, there was no regulatory pressure for the Urology Department to evaluate T2 following its adoption. This was because the reorganisation of Primary Care Trusts into Clinical Commissioning Groups had led to the disappearance of the key evaluation customers. Consequently it was unclear who was responsible for T2 in the newly formed Clinical Commissioning Groups and changes in the study Trust's internal management structure. Also unlike T1 and T3, there were no additional incentives for the Urology Department to undertake any further evaluations beyond those carried out by the developing Trust.

T1LeadConsultant expressed frustration at the lack of capacity to carry out additional evaluations that would be “clinically meaningful to patients”. This was attributed to lack of additional funding for evaluations beyond the regulatory requirements. This further hampered the T1 project team’s capacity to adequately perform the root cause analyses and follow up issues raised in the monthly VTE audits and the root cause analyses. To overcome the funding restrictions, T1LeadConsultant and T1LeadNurse often carried out the root cause analyses and other related work in their own time. These limitations created new ethical dilemmas whereby they were increasingly concerned that findings from evaluations would require appropriate follow up to ensure patient safety. These new ethical dimensions further limited the evaluations that the T1 project team was willing to undertake beyond the monthly VTE audits. The T2 and T3 project teams, especially the nurses, also raised concerns about lack of funding for evaluations. The T2 nurses and the Critical Care Outreach Team were interested in carrying out evaluations that would show decision makers within the Trust their value and contribution to their respective departments. Another barrier was that the three CDSSs were used as agents of change to restructure workflow and reassign responsibilities. There was a desire by the respective project leaders to standardise clinical processes and user practices using the CDSSs. The project teams also intended to use the CDSSs to improve clinical processes, user performance, efficiency and compliance with guidelines. For all three CDSSs, there was a presumption by project leaders that the CDSSs would automatically result in improvements in patient outcomes. However, these assumptions and their impact on clinical workflow, professional dynamics and the wider organisation were not evaluated.

Organisational barriers identified in this sub-section primarily related to the following:

- shifting organisational priorities led to changes in evaluation focus
- external pressures dictated the range and boundaries of evaluations
- top down approach to evaluations that were carried out
- pressure on evaluators to carry out only the evaluations which would be considered meaningful by key decision makers or funders of evaluations
- lack of funding or other incentives to evaluate beyond the immediate regulatory requirements
- lack of capacity or authority to effectively implement changes in response to findings from evaluations
- lack of clear customers for some evaluations
- lack of ownership of evaluations

## **8.7 Benefits of evaluations**

For all three cases, the main benefit of evaluation was to secure the continued funding of the respective CDSS projects. For T1, the monthly VTE compliance audits were directly linked to the operational funding for the study Trust from the commissioners. As such, there was a financial incentive to carry out this evaluation. The main benefit for T2 evaluations was to satisfy the regulatory requirements for its registration with the MHRA and for CE marking. This exercise was important for T2 developers to show potential commissioners that the CDSS had been robustly evaluated and that it was safe to use. Indeed, T2LeadConsultant did not see the need to commission any further evaluations because he was satisfied with the evaluations that

had been carried out at the developing Trust. The main benefits of the formal evaluations that were carried out for T3 were to produce information that would show that the expected benefits as sold by the developers had been achieved. Producing this information was important to justify the technology investment to key stakeholders and also the continued funding of T3 at the study Trust when the tender was due to be renewed. It would appear that the espoused benefits of the main studies for all three cases were primarily used for decision-making within the study Trust. The primary benefits of CDSS evaluations were to show adherence to NICE guidelines and to secure financial incentives for the study Trust. It would appear that these benefits might have been achieved at the expense of carrying out other evaluations, which would have looked at the wider CDSS effects on patient outcomes, the organisation and users. It would appear that evaluators primarily focused on retrospectively justifying investment in CDSSs by highlighting wider benefits even where there was no direct correlation with CDSSs.

## **8.8 Further factors of evaluations identified in the cases**

The three case studies revealed additional factors of evaluation that were not identified in the literature review and the resultant CDSS evaluation framework. These factors included the following:

- funding for evaluations
- stakeholders involved with the evaluations
- the criteria used to judge the success or failure of evaluations

These factors will now be discussed in turn.



### 8.8.1 Funding for evaluations

Sources of funding for CDSS evaluations ranged from the study Trust to external stakeholders and some individual CDSS project team members. Table 8.6 shows the range of resources that were allocated to each CDSS project team before, during and after implementation. All three CDSSs were reported to have been allocated adequate funding for adoption and implementation. However, it was unclear how much funding was allocated to evaluations either in absolute terms or as percentages of individual projects. The T1 project team benefited from executive support and funding from the study Trust and the Strategic Health Authority for its initial development and related clinical and technical efficacy evaluations. T1 evaluations received additional funding in the sense that key project members such as T1LeadConsultant, T1ImplementationNurse and T1Developer were fully seconded to the T1 project during its development and implementation phases. Additionally, many projects were put on hold to ensure that the study Trust's ICT Department focused primarily on T1. The Strategic Health Authority also commissioned a leading accounting firm to carry out an audit of the study Trust's VTE management system. However, subsequent evaluations such as the root cause analyses did not receive additional funding from the study Trust and were instead funded from departmental resources and evaluators' own time. The production and reporting of monthly VTE audits following T1 implementation was assigned to the study Trust's Information Department. T2 was also reported to have received adequate funding for its implementation from the Primary Care Trust and other project partners. No funding was allocated for T2 evaluations by the study Trust and the T2 project did not receive any support from the ICT Department. A medical school that was linked with the study Trust funded the T2 patient satisfaction survey. The T3 project team received executive support and funding by the study

Trust for its adoption and implementation. Additionally, because the study Trust was a flagship site for T3 development, it benefited from ongoing support from the developers who were keen to expand their commercial interests by rolling it out in the study Trust.

CDSS	Pre-implementation resources	Resources during implementation	Post implementation resources
T1	<ul style="list-style-type: none"> <li>• Secondment of key personnel to the T1 project team</li> <li>• Prioritisation of T1 project with the ICT department</li> <li>• Funding for technical and clinical efficacy testing and validation of software algorithms</li> <li>• Funding for information department to produce monthly CQUIN audit reports</li> <li>• Funding for peer reviews of guidelines and algorithms</li> </ul>	<ul style="list-style-type: none"> <li>• Evaluations carried out during implementation fully funded by the study Trust</li> <li>• Formal and informal evaluations involved multiple professional groups and clinical departments</li> <li>• T1 project team fully funded during implementation</li> </ul>	<ul style="list-style-type: none"> <li>• Information Department funded to carry out monthly CQUIN VTE risk assessment audits</li> <li>• Funding allocated to Haematology department for T1 Lead Consultant and T1 Lead Nurse</li> <li>• No funding for root cause analyses</li> <li>• No funding for VTE Link Nurse and Audit and Effectiveness audit</li> </ul>
T2	<ul style="list-style-type: none"> <li>• No formal evaluations at study Trust prior to adoption</li> <li>• Professional judgment used to support adoption decision</li> <li>• Prior knowledge of developers used to support adoption decision</li> <li>• Evaluations at developing Trust used to support adoption decision</li> </ul>	<ul style="list-style-type: none"> <li>• No formal evaluations during implementation</li> <li>• No implementation support from ICT department</li> </ul>	<ul style="list-style-type: none"> <li>• No funding allocated for formal evaluations by the study Trust</li> <li>• Patient survey funded by local medical school</li> </ul>
T3	<ul style="list-style-type: none"> <li>• No formal evaluations carried out by study Trust prior to adoption</li> </ul>	<ul style="list-style-type: none"> <li>• Evaluations during the pilot and Trust-wide implementation fully</li> </ul>	<ul style="list-style-type: none"> <li>• Benefits studies funded by study Trust</li> </ul>

	<ul style="list-style-type: none"> <li>• Professional judgement used to support adoption decision</li> <li>• Prior knowledge of developers used to support adoption decision</li> <li>• Evaluations at developing Trust used to support adoption decision</li> </ul>	<p>funded by the study Trust and T3 developers</p> <ul style="list-style-type: none"> <li>• Implementation teams fully funded, including secondment of key personnel and prioritisation of T3 project with ICT department</li> </ul>	<ul style="list-style-type: none"> <li>• No additional funding for formal evaluations</li> <li>• No funding allocated for the Critical Care Outreach team service evaluation</li> </ul>
--	--	--	---

Table 8.6 Resources allocated for CDSS project teams

The VTE Link Nurse-led audit also did not receive direct funding from the study Trust and was completed in the evaluators' own time. However, this audit widened the scope of evaluative activity and brought new dimensions from new evaluators outside the initial T1 project team. This audit unearthed many issues that were missed by the monthly CQUIN VTE audits. However, challenges arose when it came to acting on the findings of the audit because the evaluators lacked the resources and unlike the CDSS project teams they had no power to enforce the necessary changes in practice. The support and funding allocated to the T1 and T3 project teams enabled them in part to undertake some of the evaluations, particularly the technical aspects during the development and implementation stages and continuing audits related to national guidelines in the case of T1. The VTE audits were important for the study Trust because they were directly linked to financial incentives agreed with Commissioners under the CQUIN framework for VTE management. For T1 and T3, there were also financial penalties for the study Trust if the national targets were not adhered to. The potential for financial penalties due to non-adherence with these targets resulted in significant resources being allocated to the T1 and T3 Project teams to undertake the necessary evaluations to show compliance. Overall, it would appear that adequate resources for evaluations were allocated during the initial adoption and implementation stages to show CDSS technical and clinical efficacy as well as patient safety. It would appear that once the CDSSs were deemed safe for use within the defined patient groups, there were assumptions by the key stakeholders that the stated objectives would be achieved without need for any further evaluations. Post implementation evaluations appear to have been primarily focused on proving the study Trust's guideline compliance.

This sub-section has shown the importance of ensuring that CDSS evaluations are adequately funded. CDSS evaluators should consider the following:

- CDSS evaluations require adequate funding at every phase of CDSS lifecycle
- funding of CDSS evaluations should be considered as part of the adoption process
- sources of CDSS funding can be internal or external
- funders of evaluations may influence how evaluations are carried out and reported
- evaluations that are not commissioned or supported by key decision makers are likely to result in minimal effect on decision making

### **8.8.2 Characteristics of the key stakeholders involved with evaluations**

For each of the CDSSs, the owners of the evaluations were primarily the Lead Consultants who had championed their adoption. Table 8.7 shows the different evaluators and targeted audiences of CDSS evaluations. The main T1 customers were the Trust Board, the Thrombosis Committee, Commissioners and the Strategic Health Authority, whose responsibility it was to implement the government's initiatives in NHS Trusts. The primary customers of T3 evaluations were the study Trust Board, the Strategic Health Authority and Commissioners. In contrast, the circumstances surrounding T2's adoption meant that the original evaluation customers were no longer available. This was caused by the reorganisation of the Primary Care Trusts into Clinical Commissioning Groups. For T1 and T3 the main customers of evaluation dictated the course of evaluations by setting guidelines on how audit data would be reported. However, it was up to the respective Lead Consultants to decide the best evaluation methods to meet the nationally mandated requirements.

CDSS	Evaluators	Target audiences
<b>T1</b>	<ul style="list-style-type: none"> <li>• Senior Haematologists</li> <li>• Senior clinicians in the VTE Project team</li> <li>• ICT technicians and project managers</li> <li>• VTE Project team</li> <li>• Pharmacy and Haematology Departments</li> <li>• PricewaterhouseCoopers auditors</li> <li>• VTE T1LeadConsultant and Lead Nurse</li> <li>• Project Lead Nurse, VTE link nurses and Audit and Effectiveness Department facilitator</li> </ul>	<ul style="list-style-type: none"> <li>• Thrombosis Committee, Trust Board, Commissioners and Regulators</li> <li>• VTE Project team, Thrombosis Committee and Trust Board</li> <li>• VTE Project team and clinical users</li> <li>• VTE Project team, Thrombosis Committee, Trust Board, Commissioners and Regulators</li> <li>• Departmental business managers and clinical leaders</li> <li>• Trust Board, Commissioners and the Strategic Health Authority</li> <li>• Link Nurse Network, VTE Project team and the Audit and Effectiveness Department</li> </ul>
<b>T2</b>	<ul style="list-style-type: none"> <li>• T2 developers and clinical researchers</li> <li>• Clinical researchers and system's developers</li> <li>• Lead Consultant at study Trust (T2LeadConsultant)</li> <li>• T2 developers</li> <li>• Clinical researchers</li> </ul>	<ul style="list-style-type: none"> <li>• MHRA and Commissioners</li> <li>• T2 developers</li> <li>• Commissioners and Trust Board</li> <li>• MHRA and commissioners</li> <li>• Researchers and commissioners</li> </ul>
<b>T3</b>	<ul style="list-style-type: none"> <li>• T3 implementation team and line managers</li> <li>• Project Lead Consultant and Outreach Lead Nurse</li> <li>• Critical Care Department and the IQNARC Case Mix Team</li> <li>• Critical Care Department and Dr Foster</li> <li>• Critical Care Outreach Team</li> <li>• T3 Lead Consultant</li> <li>• The implementation team and T3 users</li> </ul>	<ul style="list-style-type: none"> <li>• T3 Steering Committee, Trust Board and Commissioners</li> <li>• T3 Steering Committee, Trust Board and Commissioners</li> <li>• Trust Clinical Governance team, Trust Board, Commissioners and regulators</li> <li>• Critical Care Department</li> <li>• Clinicians, Trust Clinical Governance team, Trust Board and regulators</li> <li>• T3 users, line managers, implementation team, T3 steering Committee and Trust Board</li> </ul>

Table 8.7 CDSS evaluators and target audiences

The interactions between the key stakeholders varied for the three cases. In some instances, such interactions resulted in the users' inadvertent responses to the presence of the evaluators, which were taken as positive effects of the CDSSs. For all three CDSSs, users repeatedly reported having to "cover their backs", whereby they were aware that the CDSSs would be remotely monitored by the respective project teams. In response, they made sure that they "did enough" to satisfy their legal and professional obligations. For example, the T3 implementation team worked closely with ward staff to offer them training and support during the two-week implementation phases. They informally evaluated and monitored the implementation process, particularly how ward nurses were collecting and recording clinical observations. The T3 implementation team concluded that training and implementation had been successful in areas where T3 had been implemented. However, once the implementation team left the ward areas, the culture around the recording of clinical observations was reported to have changed. Instead of recording clinical observations at the point of care as required, users were reported to be writing the observations on pieces of paper, and then inputting them all at once afterwards. In such cases, the date and time stamp would be wrong and potentially deteriorating patients also placed at risk. In part, the reason for this change was that the T3 PDA batteries were running out halfway through the cycle of clinical observation taking and also that nurses found it easier to collect observations and record them all on T3 afterwards. However, formal evaluations that were carried out did not uncover these practices and it was only highlighted by the Critical Care Outreach team during their routine follow up sessions.

This sub-section has shown that the characteristics of key stakeholders affect CDSS evaluations. The following factors need to be considered:



- it is important to gain “buy in” from key stakeholders
- the targeted audience of evaluations should be clearly defined
- using CDSSs to change user behaviour may result in resistance
- evaluations should be geared to assess interactions between stakeholders, including evaluators and CDSS users
- Hawthorne effect needs to be considered when carrying out evaluations

### **8.8.3 Criteria used to judge the success or failure of evaluations**

For all three cases, the main criteria for judging the success of CDSS evaluations were their ability to meet the project objectives. None of the evaluations were deemed to have failed. In the case of T1 and T3, the key criteria for judging success of evaluations were primarily the achievement of national targets. While there was awareness amongst the project leaders of the importance of organisational and other related issues, their evaluation priorities were bound to the CDSS adoption drivers. These drivers were CQUIN targets for VTE; the reduction of waiting times in accident and emergency for T3; and maintaining the follow up visits for stable prostate cancer patients for T2. Some key stakeholders, for example, T1ImplementationNurse and the Critical Care Outreach argued that the successes of their respective CDSSs were not well publicised within the study Trust. In all three cases, just undertaking the internal audits and informal evaluations was deemed a success given the resource limitations. For T2, its successful implementation and potential benefits ascertained from informal evaluations were considered as successes. It appears that the T3 benefits studies were clearly defined, focusing on whether the study Trust had achieved the expected benefits of adoption based on the outcomes from the developing Trust. Although it was generally accepted by key stakeholders that such benefits had

overwhelmingly been proven, it was difficult to judge these claims due to unavailability of evaluation documentation. Furthermore, there were several other Trustwide service improvement initiatives running simultaneously with T3 that were not considered in the evaluations. It is quite likely that some effects of these concurrent initiatives overlapped with T3 but may have been attributed to T3 instead. Although there was awareness amongst the T3 project leaders that these benefits may have overlapped, they did not believe that this diminished the benefits of T3 in any way. It appeared that all three project teams were keen to show the benefits of the CDSSs to the key decision makers both internally and externally so that they would gain support and continued funding. However, in many cases, there were no direct causal relationships between the reported outcomes and clinical practice. It was also unclear which evaluations had resulted in resource allocation decisions within the study Trust. This was particularly so in relation to external decisions.

This sub-section has shown that the criteria to judge the success or effectiveness of CDSS evaluations should go beyond the stated objectives of CDSS projects. The following should be considered:

- CDSS evaluations should assess various outcomes, including patient, organisational, social and users outcomes
- caution should be taken to avoid just focusing on satisfying the drivers for CDSS adoption
- successful implementation of CDSSs should not be used as the only parameter to judge successful evaluations

- efforts should be made to separate the effects of concurrent service improvement initiatives from those of CDSSs
- efforts should be made to establish causal relationships between reported outcomes and CDSSs

## **8.9 Limitations of the evaluations carried out**

All three CDSSs were driven by NICE guidelines and other related national priorities at the time. As such, all evaluations were related to the guidelines in order to be seen as credible and to elicit the support of key decision makers and funders. The evaluations undertaken for all three CDSSs missed some important issues. Although the study Trust was reported to have adhered to NICE guidelines and met the national targets for the respective clinical issues, the evaluations did not address contextual, methodological and user specific issues that would have a bearing on the primary evaluations. For example, the criteria for judging the success of the CDSSs were based on untested assumptions by key stakeholders regarding guideline adherence and perceived improvements in workflow and patient outcomes. Only a few evaluators acknowledged the limitations of the evaluations that they had carried out. T1LeadConsultant noted that they had evaluated as best as they could within the limited resources and in the absence of comparable systems in the UK at the time. T2LeadConsultant also showed awareness of the limitations of the evaluations that had been carried out but did not feel that any further evaluations were necessary at the time. However, T2Nurse1 did not feel that the evaluations that had been carried out for T2 addressed the additional challenges that were brought by its introduction. Similarly, the Critical Care Outreach Team also admitted that most of their evaluations were based on

incidental and anecdotal evidence. However, they were confident that many benefits had been achieved although they had not been formally evaluated and could not be proven by actual evidence.

All three CDSSs introduced unanticipated and unwanted problems which led to resistance in clinical areas and expected users finding ways to “get round” them. This led to many unintended uses. However, the evaluations performed did not address these problems. Furthermore, the evaluations undertaken were not geared to pick up unintended consequences of both adoption and evaluation. Some of these consequences resulted in disincentives. For example, following an audit by the Strategic Health Authority, the study Trust was ranked lower than some regional Trusts which were known to be poorly collecting and reporting CQUIN VTE risk assessment data. Although this was informally identified as a consequence, it also highlighted the shortcomings of using audits as a methodology to compare and rank different NHS Trusts. Another key issue that was missed by the CDSS evaluations was the effectiveness of the dissemination of evaluations to key stakeholders. This was made worse by the lack of project and evaluation documentation regarding the inputs of evaluations, such as funding and other resources made available, methodologies applied and detailed outcomes of evaluations. Most evaluations were not documented at all and some were based on anecdotal evidence and professional judgements and observations in the course of CDSS implementation or routine work. Additionally, documentation of the implementation and evaluation processes was also inadequate for T1 and T3, and unavailable for T2. However, the key stakeholders did not seem concerned about the lack of project and evaluation documentation. The evaluations could have benefited from interactions between evaluators, the respective CDSS project teams and targeted users to allow appropriate utilisation of evaluation information to

improve the evaluation of similar projects in the future. The evaluators could have looked at the wider effects of CDSSs on users and how the intended users interacted with the new technologies. They could also have considered the CDSSs' interaction and integration with existing legacy systems as well as non-electronic systems. However, the improved availability of information often led to further operational problems, such as lack of capacity to implement action plans and interventions to improve failing areas. Other key issues that were not evaluated include the training, implementation and evaluation documentation processes. These issues were important because all three CDSSs were being used for the first time in the study Trust.

The evaluations that were carried out for all three CDSSs had many limitations. To counter these limitations, CDSS evaluators need to consider the following:

- CDSS evaluations should look beyond guideline adherence and securing financial incentives from commissioners
- CDSS evaluations should look at wider effects on outcomes relating to patients, the wider organisation, clinical workflow and users
- evaluators should consider the limitations of their evaluations
- criteria to judge the success of CDSS evaluations should look beyond the stated project objectives
- assumptions about CDSS effectiveness should be based on clear evidence
- unanticipated and unwanted CDSS outcomes should be assessed and acknowledged
- unintended CDSS uses should be evaluated
- disincentives of evaluations should be assessed
- evaluations should be fully documented and disseminated to key stakeholders
- CDSS integration with clinical workflow and existing legacy systems should be evaluated

- Improved availability of management information through CDSSs does not necessarily translate to improved outcomes or successful evaluations
- NHS Trusts should put in place mechanisms to ensure that there is capacity to appropriately respond the evaluation findings

## **8.10 Overview of CDSS evaluations that were carried out**

This chapter has looked at the evaluations that were carried out across all three cases in relation to the key factors of evaluation that were identified in the literature review and the ensuing CDSS evaluation framework. These factors of evaluation include the purposes and context of evaluations, approaches and methods of evaluations as well as the barriers and benefits of evaluations. Other evaluation factors that were identified in the three cases include funding for evaluations, the attitudes of the key stakeholders towards evaluations, the criteria used to judge the success of evaluations and impact of evaluations on decision-making within the study Trust. The limitations of evaluations that were carried out were also discussed. The next chapter is going to focus on the how the key factors of evaluations that were discussed in this chapter relate to the wider CDSS evaluation literature and to the research questions. This discussion will also help to show how this research contributes to the CDSS literature in relation to the key evaluation factors that were identified in this research.

# Chapter 9 Discussion and Conclusions

## Introduction

This chapter discusses the implications of this research project and draws conclusions from the research findings. Section 9.1 revisits the research questions that were raised following the literature review. An overview of what was known about CDSS evaluation before this research was carried out is outlined, including some of the key research studies. An outline of what was done to answer the research questions is also given. The research findings are discussed in Section 9.2. This discussion is based on the CDSS evaluation framework that was developed from the literature review (see Figure 9.1). Section 9.3 presents the research findings in the context of the CDSS evaluation framework, noting the expected, conflicting and unexpected findings. Section 9.4 outlines the key research findings, and the evaluation factors that extend the CDSS evaluation framework discussed in 9.5. Section 9.6 discusses the key contributions of this research to CDSS evaluation literature, and potential limitations noted in 9.7. In Section 9.8, the recommendations for further research and conclusions are given.

## 9.1 Research questions

This research project sought to make contributions to CDSS evaluation literature across all the five key factors identified in the CDSS evaluation framework (see Figure 9.1). Using this framework, this research sought to answer the following questions:

- 1 What are the key factors that affect CDSS evaluations in a typical NHS hospital setting?
- 2 How do these factors relate to the CDSS evaluation framework that was developed from the CDSS literature review?

- 3 To what extent do evaluations affect decisions to adopt CDSSs in healthcare settings and which evaluation methods are most likely to inform CDSS adoption decisions and why?



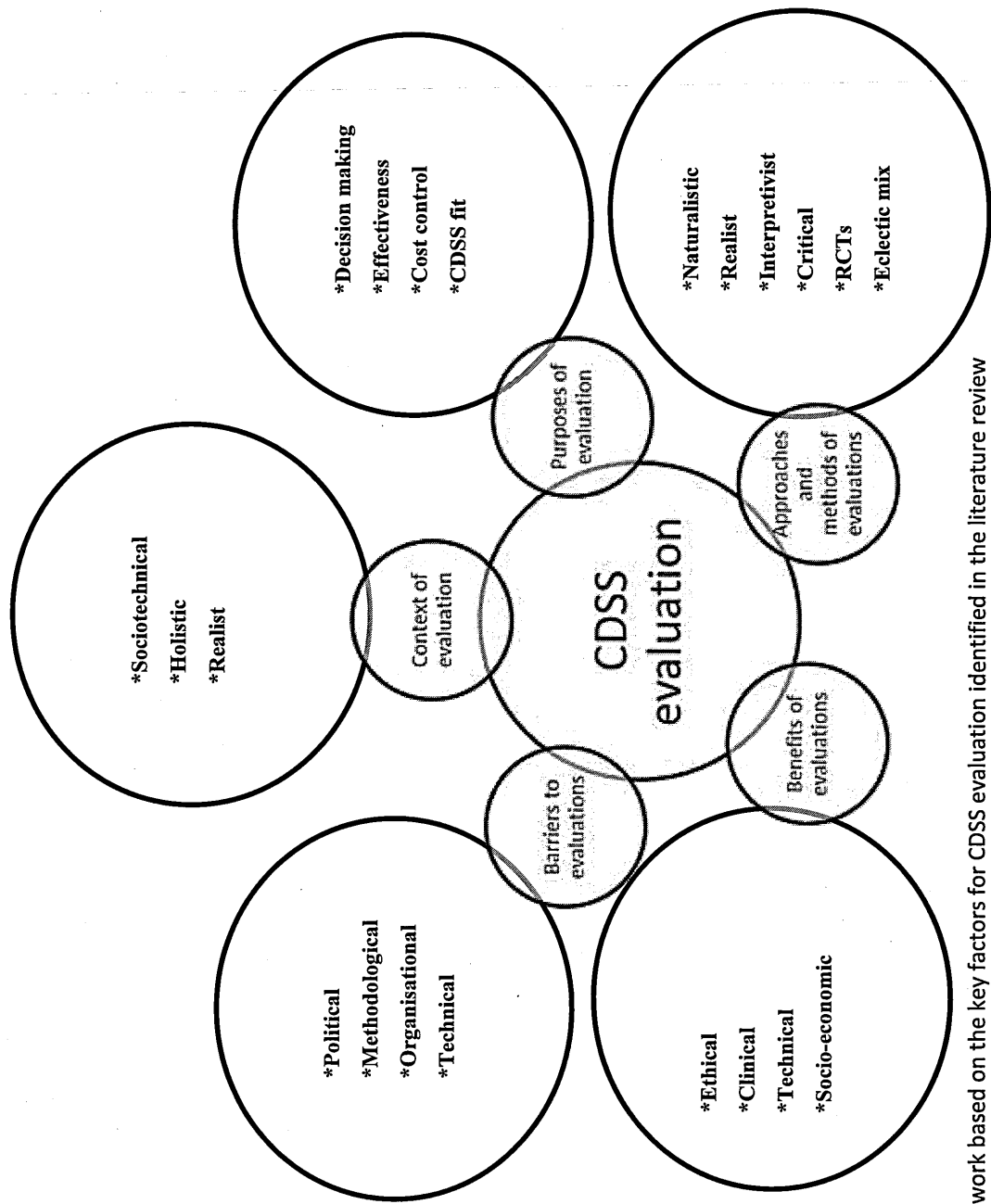


Figure 2.2 CDSS evaluation framework based on the key factors for CDSS evaluation identified in the literature review

## 9.2 Research results

The five key factors of CDSS evaluations noted in CDSS evaluation framework (Figure 9.1) were evident to varying extents in all three case studies. Some evaluation factors were more relevant to individual CDSSs depending on the circumstances around their adoption. These factors are now discussed in turn.

### 9.2.1 Contextual and organisational and social issues that affect CDSS evaluations

Since the early 2000s, contextual, organisational and social factors have increasingly been identified as important to any CDSS evaluation activity. Interviews that were carried out with key stakeholders across all three cases revealed a general consensus that contextual and organisational issues were considered important. The interviews also revealed that the key stakeholders perceived different CDSS evaluation contexts as shown in Figure 9.2.

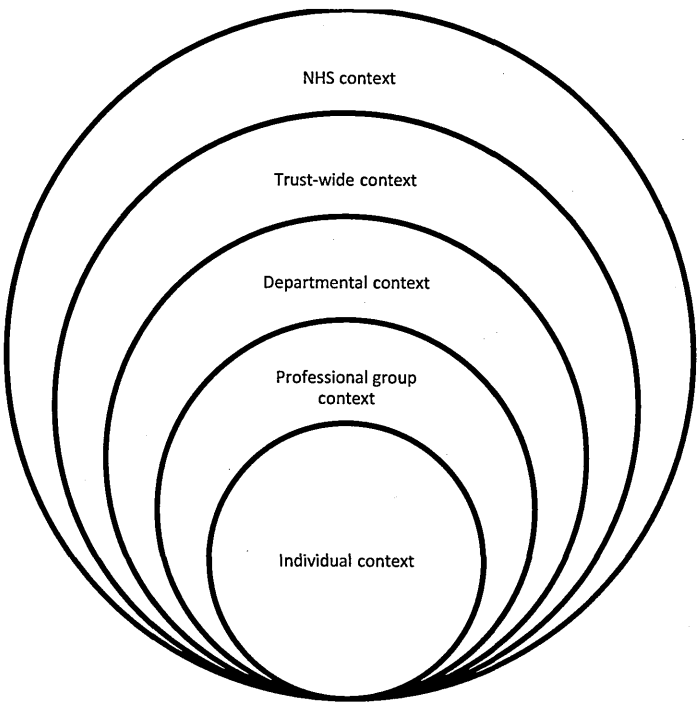


Figure 9.2 CDSS evaluation contexts as perceived by the key stakeholders

The different contexts of CDSS evaluations that are shown in Figure 9.2 are now discussed in turn. There were assumptions amongst the key stakeholders, especially the respective CDSS project leaders and senior managers that the NHS was homogenous and that each individual NHS Trust faced similar issues with respect to CDSSs and wider clinical services which they supported. The level of standardisation of healthcare systems reinforced these views across the NHS through regulatory mechanisms such as the NICE guidelines and other centrally driven initiatives, with support from professional and collegiate organisations such as the Royal Colleges of Medicine among others. These assumptions also extended to the evaluation of CDSSs in the study Trust. However, the three cases revealed that although the study Trust itself was an important context, it was not homogenous and the three CDSSs in the cases operated in very specific local conditions. This lack of a homogenous context added to the complexity of the CDSSs' evaluation. Although the key stakeholders identified the organisational context as important, little attention was paid to the contextual homogeneity in any of the CDSS evaluations that were carried out. There were a number of dimensions that affected how the local conditions within the context changed.

To begin with, these dimensions were:

- differences in existing IT infrastructure between different clinical specialties
- differences in stakeholders' attitudes towards the CDSSs and their evaluations
- stakeholders' attitudes to NICE guidelines and related clinical pathways
- existing professional and organisational culture
- motivation to comply with national guidelines and recommendations from professional and collegiate bodies
- stakeholders' ethical, professional and legal accountability

The introduction of T1 and T3 brought new powers to the Haematology and Critical Care Departments respectively. The individual project leaders also had the backing of the study

Trust's CEO and Medical Director amongst other key stakeholders and decision-makers. This enabled them to implement interventions across the study Trust and enforce compliance, which had previously failed for many years. According to the project leaders, some clinical departments that had previously failed or refused to comply with national guidelines suddenly had no choice but to cooperate with the T1 and T3 project teams. The main consequence to mitigate these powers was the tendency by some CDSS users to "tick the box" and "cover their backs". This meant that they did just enough to be seen to be compliant even though they did not believe that this would result in any improvements in their work processes or patient outcomes. This also applied to senior clinicians and managers who were keen to be seen to be complying with national guidelines and show their accountability professionally, ethically and legally to the key stakeholders such as the Trust Board, the Commissioners as well as professional and regulatory bodies.

There was a tendency by some of the key stakeholders to take both internal and external CDSS evaluations at face value across all three cases. For example, T2 and T3 project leaders trusted their peers who had developed and evaluated these CDSSs at the respective developing Trusts and did not question the evaluations that had been carried out prior to their adoption in the study Trust. This may be because all three project leaders were physicians and like most of their peers in other NHS Trusts, tend to favour quantitative over qualitative methods because they are perceived to be bias free and evidence-based. It could also be argued that where evaluations have already been carried out externally, it may be perceived as wastage of resources to carry out further evaluations internally. However, external evaluations should be interpreted cautiously in part because the contexts of evaluation are different and also because these evaluations may not necessarily reflect or serve the best interests of the adopting organisation. The evaluations that were

carried out for T2 and T3 at the respective developing Trusts addressed single outcome measures that assessed changes in the care management system, user performance and perceived effects on patient care within the settings where they were developed, rather than the wider organisational effects of these CDSSs. Additional evaluations needed to be carried out by the study trust to look at the CDSSs' wider effects on the organisation and local clinical contexts. There was also a notion by some senior clinicians that "we cannot evaluate everything" because there were limited resources to carry out the evaluations and also that evaluations would be detrimental to CDSS adoption efforts. Other key stakeholders (for example T1LeadNurse and T1LeadConsultant) were also concerned that they did not have adequate resources to deal with findings from CDSS evaluations and therefore it "would be pointless" to carry out any additional evaluations. Consequently, this raised new ethical questions regarding how much evaluation they were willing to carry out given the limited resources that were allocated for evaluations and the subsequent follow-up which they felt was required to make the evaluations worthwhile. There was also a general consensus by most of the key stakeholders that the three CDSSs were required by the study Trust regardless of whether their benefits had been proven or not. This belief was based on the assumption by the key stakeholders that by adopting the CDSSs, the study Trust had resolved many existing problems thus rendering any evaluations irrelevant. In all three cases, the project leaders saw opportunities to acquire investments in CDSSs whose clinical benefits may not have been immediately apparent at the time but would eventually be realised in the future. Expected future uses of the CDSSs included the ready availability of audit trails and performance data that could be used to monitor the clinical areas. All three project leaders were also keen to build patient databases for future clinical research projects.

This section has shown the dilemma faced by CDSS adopters regarding whether they accept external evaluations or carry out their own evaluations before implementation. By taking external evaluations at “face value”, they risked adopting CDSSs which may not be suitable for their local clinical environments. On one hand, it could be argued that even where external evaluations had been carried out, it may still be necessary to carry out internal evaluations to assess their suitability and potential effects on the local clinical settings before implementation. On the other hand, it could also be argued that limited resources could be wasted by repeating evaluations which had already been carried out externally. The onus is therefore on local decision makers to systematically evaluate external evaluations to establish their validity and also to ensure that they provide answers to questions which they may have. This is particularly important because as shown in Figure 9.2, the perceived contexts of evaluations may be different depending on the stakeholders involved. External evaluations that may be satisfactory for one professional group may be considered inadequate for another group.

### **9.2.2 Purposes of CDSS evaluations**

The purposes of CDSS evaluations were affected by key factors such as the attitudes of the key stakeholders, motivations of the evaluators and related organisational issues. The majority of the CDSS evaluations were driven by the need to show the study Trust’s compliance with the respective NICE guidelines and the national priorities at the time. It appeared that the evaluators were keen to show that they had demonstrated that the clinical interventions, which were supported by the CDSSs, had adhered to NICE guidelines. Guideline adherence was seen as a way to show the credibility, technical efficacy and safety of the CDSSs and also justify the technology investment. Aligning CDSS evaluations to the

respective NICE guidelines and other important issues at the time was also done to secure funding and gain support from key decision makers in the study Trust and commissioners who were under constant pressure to respond to national initiatives and show compliance with guidelines. The main challenge with aligning evaluation purposes to national issues was that the motivations for evaluations were defined by national guidelines and current priorities at the time rather than issues that may have been considered important at organisational or other contextual levels. Some key stakeholders were concerned that once another issue became important at national level, decision-makers at the study Trust would shift their interest in guideline adherence and related evaluations accordingly. This was particularly the case for T1 and T3 because funding for the respective projects was directly linked with meeting national targets. However, by primarily targeting the evaluations to commissioners and regulatory authorities, the project teams inadvertently alienated the users of the CDSSs. Although some evaluations of T2 (at the developing Trust) related to NICE guidelines, there was no pressure for the T2 project team at the study Trust to show that national targets were being met. However, the Uro-oncology nurses who expected to use T2 had not been involved in the adoption decision or any evaluations that had been carried out and thus lacked ownership for these evaluations.

The attitudes of the different professional groups and their disparate departments had significant repercussions for the CDSS evaluations that were carried out. T1 and T3 were implemented across the study Trust with little regard to the differences between the clinical specialties which they supported. There appeared to have been general assumptions by some of the key stakeholders that T1 and T3 would fit any department in the study Trust and that the intended users would appreciate the CDSSs' importance because they were both linked to NICE guidelines. However, there were significant

variations in existing computerised legacy system infrastructures between the different departments. Additionally, the key stakeholders had diverse attitudes towards the value of these CDSSs. Their value sets around CDSS evaluation were also confined to their respective departments and in some cases, their professional and collegiate affiliations. Some clinical areas did not feel that they needed these CDSSs and did not see any value in their evaluation. For example, highly specialised areas such as the cardiology and neurology departments felt that their existing computerised legacy systems were more specialised and better suited to their clinical environments than T3. They also felt that their existing clinical pathways adequately covered the new NICE VTE guidelines which T1 was based on and therefore their respective departments did not require the additional decision support. To these departments, T1 evaluations, such as the CQUIN VTE audits were viewed as administrative exercises that did not fully report the extent of their existing VTE management efforts. There were also important departmental differences between specialised and high demand areas, which were often in the spotlight regarding national issues compared with general medical departments. By virtue of their experiences dealing with these issues, these departments felt that they were better equipped to make their own decisions to suit their respective clinical areas without involving the CDSS project teams. However, their capacity to effectively implement and adhere to these guidelines had not been evaluated. It could also be argued that allowing individual clinical departments to autonomously implement national guidelines and subsequently evaluate their own adherence would potentially result in variations in guideline adherence, clinical practice and ultimately, patient outcomes. Indeed, the clinical departments were likely to have different priorities and motivations for evaluation which would have affected the wider organisational goals. Furthermore, not all the clinical departments had the



experience, expertise or the necessary resources to carry out such implementation and the required evaluations.

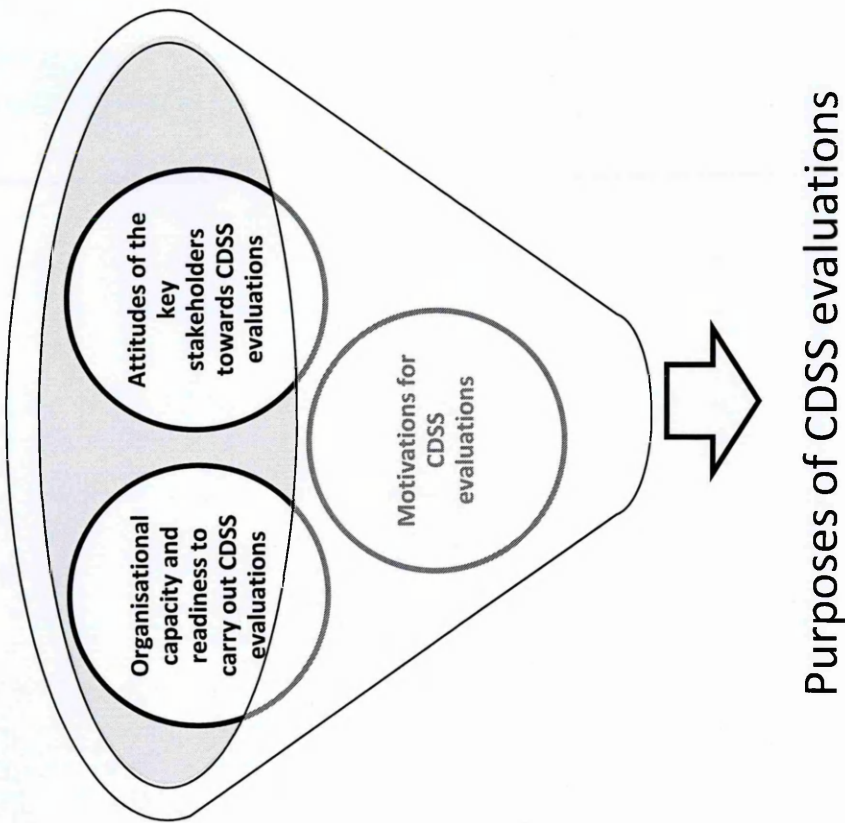
The evaluators for all three CDSSs showed a desire to improve clinical processes and patient outcomes. However, it could be argued that some of the evaluations that were carried out had no causal links to the respective CDSSs as reported, although the evaluators sought to emphasise such links. For example, the root cause analyses investigations were reported to have improved as a result of T1 and T3 introduction. However, the root cause analysis investigations themselves were not evaluations of T1 or T3 but instead looked at failures in the wider clinical processes that were supported by these CDSSs. However, the root cause analysis investigations exposed some of the shortcomings of the primary CDSS evaluations. For example, some patients who had been deemed to have had valid VTE risk assessments had not been prescribed the thromboprophylaxis as recommended by the NICE guidelines and where prescribed, in some cases the nurses did not administer thromboprophylaxis. T1 and T3 project leaders also reported that the processes of carrying out routine audits had significantly improved following the successful implementation of both CDSSs. However, the improvements in these processes were not in themselves improvements in patient outcomes as claimed by the evaluators.

By aligning the CDSS evaluations to NICE guidelines and other national initiatives, there was an expectation by the decision makers that patient outcomes would improve by default. Nevertheless, none of the evaluations focused on specific patient outcomes. The respective CDSS project team leaders were also keen to show that introduction of the CDSSs had resulted in the improvement of their overall clinical management systems. This premise was primarily based on the improved accessibility of previously unavailable management information. This information enabled CDSS project leaders and line managers to

undertake audits and benchmark their services against national standards. However, none of the evaluations were specifically focused on looking at the clinical management systems. Furthermore, all three project leaders and other key stakeholders revealed that the CDSSs were not being fully optimised as intended by the developers. For example, although all three CDSSs had the capability to provide various management reports and data mining opportunities, the project leaders noted that these facilities were not being utilised to learn about the CDSSs, nor help to improve clinical processes and patient outcomes. The main reason given for not taking advantage of these opportunities was the lack of time and other high priority issues that line managers and departmental leaders had to deal with simultaneously. Most line managers and some of the departmental leaders were reported to have been unaware of the availability of this information, and in some cases did not know how to access or use this information. The primary evaluations for all three CDSSs were focused on financial incentives that were attached to the services which they supported. For T1 and T3, these incentives were related to direct payments from the commissioners towards their operational budgets when they showed compliance with NICE guidelines. The main evaluations that were carried out for T2 were also fully funded by the commissioners who were keen to show compliance with NICE guideline recommendations and gain clinical, operational and financial benefits in the long run.

The purposes of CDSS evaluations should include a range of technical, clinical, organisational, human and social issues. These evaluations should take into consideration the human CDSS interaction within the context of complex hospital settings where there are multiple professional groups who exhibit different attitudes towards CDSSs and their evaluations. It is essential to evaluate the CDSS for technical and clinical efficacy as well as its adherence to the underlying guidelines. However, successful technical and clinical

evaluations do not necessarily guarantee that the CDSS will be accepted and used as intended. There is also no guarantee that guidelines will be adhered to and ultimately, patient outcomes may not be improved or maintained. It is therefore essential to widen the purposes of evaluations to cover sociotechnical aspects to understand their effects on established social systems. Figure 9.3 shows the factors which affected the purposes of CDSS evaluations for all three cases and some key questions which need to be considered when looking at the purposes of CDSS evaluations.



#### Evaluation purposes - key questions for evaluators to consider

- What are the CDSS project goals?
- Have the processes that lead to the achievement of these goals been clearly identified?
- Have all the key stakeholders that are involved in these processes been identified?
- What evaluations need to be carried out and why?
- Which stage of CDSS lifecycle needs evaluation?
- Who are the evaluation customers and what are their information needs?
- What are the ultimate decisions that will be affected by the evaluations?
- Who is going to carry out the evaluations and what are their backgrounds and motivations?
- Do they have the necessary experience and expertise to carry out the evaluations?
- Is there adequate funding to carry out evaluations?
- What incentives are available for the evaluations?
- Are there clear boundaries of the evaluations and defined parameters for success and failure?
- What methods can best be used to carry out these evaluations at this stage?
- Do the key stakeholders accept these methods and understand the need and focus of the evaluations?
- Are the intended users involved in the evaluations?
- Do these evaluations also look at unintended, undesirable and unwanted outcomes of CDSSs?

Figure 9.3 Factors affecting the purposes of CDSS evaluations and the key questions that need to be considered by the evaluators

### 9.2.3 Approaches and methods of CDSS evaluations

The literature review suggested that there was scope to use a range of evaluation methods throughout a CDSSs' lifecycle. However, positivist approaches dominate CDSS evaluation literature, thus restricting the range of methods used and hence limiting the scope of findings and recommendations for practice. The majority of the evaluations that were carried out for all three CDSSs in the cases adopted a positivist stance and focused on single outcome issues which were isolated from the real clinical environments. The methods applied were often informal before and after assessments, simple audits and peer reviews to show compliance with national guidelines and related regulatory requirements and to ensure financial stability for the study Trust. Interpretivist and critical approaches could have helped to set the wider boundaries and scope of evaluations by identifying the interrelationships between key stakeholders and the related contextual issues. There was inadequate project and evaluation documentation for all three CDSSs. Indeed, most of the evaluations were not documented at all. Some evaluations were based on anecdotal evidence as well as personal and professional judgements, and routine observations in the course of CDSS development/procurement, implementation or in the course of the routine work of the stakeholders involved. The lack of documentation may suggest lack of rigour of evaluation methods or even deliberate attempts to restrict access to evaluations in some cases.

The majority of evaluations that were carried out for all three CDSSs were formative. While formative evaluations are essential, especially for the developmental and early implementation phases of the CDSS lifecycle, summative evaluations could have provided a wider view of the real effects of the CDSSs across the care delivery system and wider

organisation. It does appear that there were no incentives for commissioning summative evaluations because decision makers at the study Trust were gaining enough information to satisfy the primary CDSS evaluation audiences from the formative studies. Although some of the key stakeholders acknowledged the methodological limitations, they were satisfied with the respective evaluations that had been carried out. Also the novelty of all three CDSSs appears to have affected the range of evaluations that could be carried out because there were limited studies of evaluations that had been carried out within the NHS context. T3 project leaders reported that they had carried out two benefits realisation studies. However, due to lack of evaluation documentation, it was difficult to assess how the benefits realisation methodology had been applied to the benefits studies that were carried out for T3. However, a key flaw of the reported benefits realisation evaluations was the lack of involvement of any stakeholders outside the core T1 project team and their peers from the Critical Care Department. Table 9.1 shows some of the key questions that evaluators need to ask to enable them to establish which evaluation methods would be most suitable to provide them with adequate answers throughout a CDSS's lifecycle.

What problem (s) is the CDSS intended to solve? How does it work? Is it integrated with existing legacy systems? What are the expected benefits of the CDSS? How do we know when we have achieved these benefits?
Does the CDSS work as intended? Does it improve work processes? Is it accepted by users? Is it being used as intended? Does it increase their job satisfaction?
What is the cost of developing or buying the CDSS? Is it worth developing or buying the CDSS? How does it compare with current system?
What are the actual and potential pitfalls of the CDSS? How much is it integrated into clinical workflow?
What are the CDSS's organisational and social consequences? How well does it fit the organisation and clinical contexts where it is introduced? What is the impact of the CDSS? Are the benefits sustainable?

Table 9.1 Key evaluation questions that could guide which CDSS evaluation methods to use

The key evaluation questions that are noted in Table 9.1 and possible evaluation methods to help to answer them are now discussed in turn. Questions relating to the problems which the CDSS is intended to solve relate primarily to the assessment of the users’ needs. This requires CDSS developers to engage with the intended CDSS users to gain a full understanding of their work processes and the context where the CDSS will be used. Indeed, the needs assessment can inform decision-makers whether the CDSS is required at all. These needs assessment can also guide decision-makers whether they develop the CDSS internally, buy off the shelf or commission external developers to develop a bespoke system. Qualitative methods such as interviews, documentary analysis and observations may be used at this stage to gain an in-depth understanding of the users’ requirements. Documentary analysis may include a review of related literature, clinical pathways and related guidelines. It may also be necessary to write paper based algorithms which take into account all aspects of the clinical guidelines that support the CDSS. These algorithms

can then be translated into software algorithms which will form the basis of the CDSS. Cost-benefit analyses will also need to be undertaken to establish whether it is worthwhile to replace the existing systems. The key evaluation audiences at this stage include the developers, intended users and those who are responsible for funding the CDSS. The next stage involves the design and development of the CDSS. This phase may be guided by clinical guidelines and pathways to ensure that the software algorithms adhere to the expected standards of care. It may also be necessary to adhere to the requirements of regulators such as the MHRA and product certification agencies. For CDSSs to be registered with the MHRA and to attain CE certification, they have to undergo clinical trials to ensure patient safety and to show that they function as intended by the developers. The key audiences for the evaluations at this stage include the regulatory authorities, certification agencies, CDSS funders and the developers. It will be necessary to carry out CDSS usability tests with real clinical users to test whether its key functions and ensure that it is usable within the intended environment. These evaluations are essential for both the developers and users. Where necessary, changes should be implemented before progressing further. Additional evaluations may also be carried out in artificial or simulated environments to test various clinical scenarios using contrived and real patient data. These tests are important, especially where the CDSS is based on guidelines, with direct effects on clinical decision making and therapeutic effects for patients. These tests enable developers to test the CDSS's adherence to the key aspects of guidelines and clinical practice. The key evaluation audiences at this stage include the CDSS developers, users, funders and the wider healthcare and research community.

Field tests of the CDSSs in the clinical environments where they are intended to be used are essential. Evaluations may be carried out at the pilot implementation stage to establish



whether the CDSS functions as intended and that it fits into the clinical workflow. Some CDSSs may require testing with different groups of professionals such as nurses, physicians and administrators to see whether they function as intended, help to carry out tasks and whether users accept and use them. Most CDSSs bring about changes which result in significant service redesign and require changes to users' behaviour. The literature review and the case studies showed that such CDSSs were likely to meet resistance from intended users. Evaluations at this stage may help to understand the patterns of CDSS usage and users' attitudes towards the CDSSs. Users may also contribute to the CDSS's evaluation, thus informing further understanding and additional developments where necessary. User training also needs to be evaluated to assess its effectiveness and sufficiency and to ensure that the CDSS is fully understood and used as intended. Where CDSS are based on existing clinical guidelines and pathways, it is essential to highlight these linkages and show how the CDSS can help users to continue or improve their adherence to these guidelines. This may also help to improve "buy in" from users if these benefits are highlighted. The key customers of these evaluations include CDSS developers, users, funders and potential purchasers. These evaluations can also be shared with the wider research community for peer reviews and further contributions.

Evaluations during the implementation stage help to assess the effects of the CDSS on the intended processes and users' behaviour. Evaluators should look at both the positive and negative effects of the CDSSs rather than just focusing on the intended outcomes. The key aspects of implementation evaluations include the CDSSs' effects on clinical workflow, quality of care, users' job satisfaction and whether it is being used as intended. Indeed, some CDSSs may produce high quality information but may be difficult to use, time consuming and poorly integrated with existing systems and clinical workflow. Evaluators

should assess structural and sociotechnical aspects to help to gain an understanding of why CDSS may or may not be accepted and used. User satisfaction also has an impact on the outcome quality such as patient care if they are not used as intended. Where CDSSs are implemented across a hospital, there may be different effects in different departments, which may also reflect the different attitudes and usage patterns. Evaluators should treat these departments as different entities rather than treating all clinical areas as a single context. The key customers of evaluations at this stage include the CDSS developers, users, funders, potential purchasers and the wider research community. There may also be regulatory requirements to report the technical and clinical evaluations to regulatory authorities.

Finally, it is essential to carry out impact evaluations on different aspects of the CDSS. This may be a series of evaluations or a summative study looking at the effects of the CDSSs in relation to the original problems which it was intended to solve. There may also be unintended consequences of the CDSSs and the evaluations themselves that may affect how the CDSS is perceived or used. For example, some evaluations may reveal negative aspects of the users, their attitudes or performance. If not handled correctly, this may result in a backlash or poor buy in. Similarly, there may be lack of acceptance and unintended uses especially where CDSSs' usage is compulsory. Summative evaluations help to establish all the operational benefits and pitfalls over time and how improvements can be implemented. Additionally, the actual CDSS costs in terms of the total investment and operational aspects can also be established. The organisation may thus be able to calculate its return on investment throughout the CDSSs' lifecycle.

This section has shown that there is scope for various methods of evaluations that can be used at different stages of a CDSSs' lifecycle. The evaluation methods should be determined

on the basis of the stages of the CDSS's lifecycle and the questions which need to be answered. These evaluations can be at CDSS level, professionals, organisational or sociotechnical levels. There is wide scope to use positivist, interpretivist and critical approaches to broaden the scope of evaluations and gain a better understanding of their effects. There is also need to involve members of the healthcare multi-disciplinary teams to gain expertise from various evaluation traditions. Such collaborations can involve academic researchers and the wider research community to reduce the research-practice gap. Figure 9.4 summarises this section and shows the focus of evaluations at the pre-implementation, implementation and post implementation stages of the CDSS lifecycle. It also shows the key questions that evaluators should consider and to guide them to focus their evaluations appropriately.



## 9.2.4 Benefits of CDSS evaluations

Many interview participants, including the project leaders, perceived that the primary benefits of the CDSS evaluations were to secure the financial stability and reputation of the study Trust. These evaluations provided a means to show compliance with national guidelines, and in return gain funding from commissioners for the CDSS projects and secure operational income for the study Trust. There was also a desire by the decision makers to change and standardise clinical processes using the CDSSs because they were based on NICE guidelines. This resulted in assumptions by the CDSS project leaders that introduction of the CDSSs would automatically result in improvements in clinical processes and patient outcomes. However, these assumptions were not tested and the wider organisational implications of each CDSS were not evaluated. There was also an assumption that any improvements that were observed following CDSS implementation were a result of the CDSS, although in most cases, there was no evidence gathered to show causal links between the CDSSs and the reported operational, financial and clinical outcomes. Indeed, none of the evaluations explored the links between the CDSSs and patient outcomes. It could be argued that given the number of simultaneous improvement initiatives that had been implemented at the same time, it was difficult to clearly establish the links and to separate the effects of other contributory variables.

T1 evaluations were reported to have helped to secure funding to develop the Haematology Department. The funding was used to create a new role for a Specialist VTE Lead Nurse (VTELeadNurse) and a new position for a Consultant Haematologist. However, other comparable NHS Trusts were reported to have invested in larger multi-disciplinary VTE management teams and created new roles for nurses to carry out the Root Cause

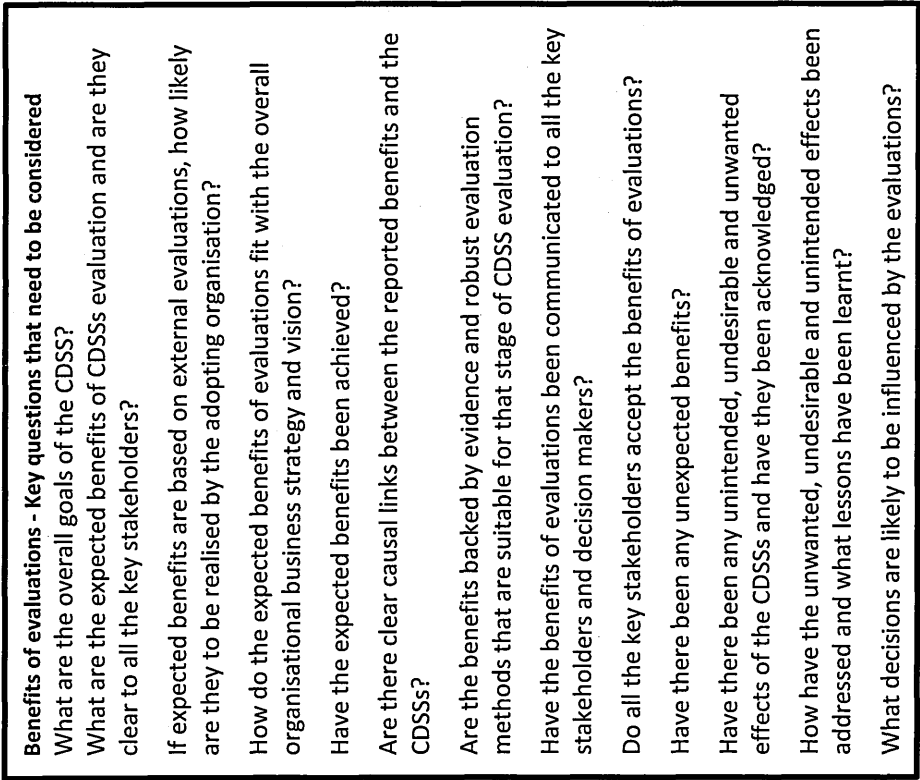
Analyses for hospital acquired VTE. T1 project leaders argued that similar investments in the study Trust could have improved their capability to carry out more evaluations and also enabled them to follow up any issues that were raised by the current evaluations. Although patient outcomes were not directly evaluated, the VTE project team reported that T1 had improved VTE awareness and clinical decision-making across the study Trust. They also noted that a significant reduction of variations in clinical practice had been observed through improved guideline adherence and easy accessibility of data. As a result, the VTE project team and line managers were able to monitor performance and the validity of VTE risk assessments that had been carried out. The improved accessibility of VTE data helped the audit and reporting processes as well as the Root Cause Analysis for VTE processes. These benefits were expected to improve the awareness and quality of VTE risk assessments and ultimately reduce the overall incidence of hospital acquired VTE. Although no formal evaluations were commissioned for T2 by the study Trust, T2LeadConsultant noted various expected benefits. To start with, T2 provided a single database to manage all patients with stable prostate cancer (from different consultants), which was reported to bring operational, financial and clinical benefits. Although these benefits had not been evaluated, T2LeadConsultant believed that there would ultimately be improvements in patient outcomes in the long term. There were also expected future uses of the patient database to identify patients who were suitable for clinical research projects, which would be beneficial for both the study Trust and the patients themselves. Furthermore, the CDSS project leaders expected that changes in clinical practice and guideline requirements would also be easily implementable by revising the software algorithms that underpinned T2. Other expected benefits that were reported to have been achieved by the T2 developing Trust, such as financial savings, upskilling inexperienced staff and integration of T2 with

existing legacy systems were not evaluated at the study Trust. Similarly for T3, it was not clear to what extent the study Trust had achieved the clinical, operational and financial benefits that had been reported by the T3 developing Trust. The lack of evaluation documentation from the two benefits realisation studies that had reportedly been carried out for T3 made it impossible to assess whether the reported benefits had been achieved and to what extent. T3 was also reported to have brought improvements to operational processes around tracking and triggering potentially deteriorating patients in comparison with the paper-based clinical observation charts. Additionally, it was reported to have enabled the upskilling of less experienced nurses to support their clinical decision-making. The easy accessibility of patient data made it easier to monitor ward areas remotely, identify failing areas and target interventions aimed at improving performance and ultimately patient outcomes through the Critical Care Outreach Team. However, these benefits had not been formally evaluated. Also, some of the reported improvements in patient outcomes were based on wider service evaluations which did not directly assess the effects of T3 on the reported outcomes. Other aspects that were not formally evaluated include the usage and acceptance; unintended usage; anticipated, unanticipated, undesirable and unwanted problems.

This section has looked at the range of benefits of CDSS evaluations that were reported in the three cases. These benefits were mainly in the form of financial incentives and adhering to the national guidelines. Some benefits were department specific rather than Trust-wide, while others related to the easy availability of performance information for line managers and clinical leaders. The majority of the benefits of evaluations were based on assumptions that were not backed by evidence. Also, there was no causal link between the reported benefits and patient outcomes. Figure 9.5 summarises the main benefits of CDSS

evaluations in the order of their priority to the evaluators of the three CDSSs. It also shows the key questions that need to be considered by evaluators regarding the benefits of CDSS evaluations.





**Benefits of evaluations - Key questions that need to be considered**

- What are the overall goals of the CDSS?
- What are the expected benefits of CDSSs evaluation and are they clear to all the key stakeholders?
- If expected benefits are based on external evaluations, how likely are they to be realised by the adopting organisation?
- How do the expected benefits of evaluations fit with the overall organisational business strategy and vision?
- Have the expected benefits been achieved?
- Are there clear causal links between the reported benefits and the CDSSs?
- Are the benefits backed by evidence and robust evaluation methods that are suitable for that stage of CDSS evaluation?
- Have the benefits of evaluations been communicated to all the key stakeholders and decision makers?
- Do all the key stakeholders accept the benefits of evaluations?
- Have there been any unexpected benefits?
- Have there been any unintended, undesirable and unwanted effects of the CDSSs and have they been acknowledged?
- How have the unwanted, undesirable and unintended effects been addressed and what lessons have been learnt?
- What decisions are likely to be influenced by the evaluations?

Figure 9.5 Hierarchy of benefits of CDSS evaluations identified in the cases and the key questions that need to be considered by evaluators

### 9.2.5 Barriers to CDSS evaluations

The following classes of barriers to CDSS evaluations were identified in the literature review:

- Technical
- Methodological
- Human
- Organisational

These barriers were evident in varying degrees for all three cases. The main barrier that was reported by the key stakeholders was the lack of funding for CDSS evaluations. However, even where funding was made available, there were also limitations in the purposes, scope and methods of evaluation. A key contributory factor was that while the national guidelines encouraged the adoption of CDSSs, they did not provide any recommendations for their evaluation. T1 project leaders highlighted barriers to effective evaluations, which they argued were caused by failure of government (through its arms' length bodies) to insist on more robust reporting of VTE audit results by NHS Trusts. They felt that the study Trust's VTE data collection and reporting methods were more robust than some of the NHS Trusts that were ranked higher than the study Trust on national league tables. They also felt that the relative view of evaluations by the Department of Health accepted individual Trusts' self-reporting on face value while inadvertently penalising the study Trust and others whose evaluations were "more robust". Another organisational barrier to evaluations was the compulsory nature of all three CDSSs, which led to resistance and unintended uses. Although on face value it appeared that the three CDSSs were being used as intended, the VTE Link Nurse-led audit revealed that users were

mainly doing just enough to be seen to be using T1 and did not value it as a CDSS or the evaluations which were being carried out. It would appear that some CDSS users may have modified their behaviours to appear as if they were using the CDSSs as intended because they were aware that they were being monitored. This is commonly referred to as the Hawthorne effect. None of the CDSSs were geared to identify when users had done just enough to “cover their backs” or enough to “just tick the box”. The changes in the structure of NHS organisations and shifts in the focus of evaluation customers also affected all three CDSSs. This was particularly so for T2 because it was commissioned by the now defunct Primary Care Trust, which evolved into three Clinical Commissioning Groups in the city where the study Trust was based. This led to the phenomenon of the “disappearing advocate”, whereby the key decision makers and customers of evaluations, who had commissioned T2, moved to new departments or to other organisations and thus the key evaluation audience was lost. Unlike T1 and T3, there was thus no pressure for the T2 project team or the study Trust to commission any evaluations once it was adopted.

The main human barrier was that some of the key stakeholders did not see the need to evaluate the CDSSs at all. For example, T2LeadConsultant and T3LeadConsultant were satisfied with the evaluations that had been carried out in the Trusts that developed the CDSSs. Their justification for not carrying out any further evaluations was that they were satisfied with how their peers had carried out the evaluations and also that there was no need to repeat these evaluations. Furthermore, T2 and T3’s registration with the MHRA and CE marking provided additional evidence of their technical and clinical efficacy. However, although these registrations provided some reassurance regarding the CDSSs efficacy and safety, they did not provide any information about their socio-technical aspects and fit into the study Trust context. Indeed, T2 was registered with the MHRA as a

CDSS that would be used in community settings by less experienced Band 5 nurses rather than in acute hospital settings by very experienced Band 7 and 8 Uro-oncology Specialist Nurses as was the case with the study Trust. It was unlikely that the study Trust achieved the financial benefits that had been achieved by the developing Trust because the Uro-oncology Specialist Nurses cost a lot more than Band 5 nurses. Furthermore, these experienced nurses may not have required the same level of decision support as that required by less experienced Band 5 nurses. However, it could also be argued that any CDSS, especially those that are based on NICE guidelines or local clinical pathways should be fool-proof, regardless of how experienced the targeted users are. Although T3 project leaders reported that they were satisfied with the evaluations carried out by the developing Trust, they found T3 “completely unusable” on pilot implementation because it was not compatible with the existing legacy systems at the study Trust. This raises concerns regarding the robustness of prior evaluations and the information that was used to support its adoption in the study Trust. It also supports the argument that evaluation contexts are different between NHS Trusts and systematic local evaluations should be carried out to assess CDSS suitability before implementation.

For some senior clinicians and clinical line managers, the introduction of these CDSSs had resolved many existing problems. Some argued that they had grappled with non-compliance with NICE guidelines and local clinical pathways for many years and the solutions that had been provided by these CDSSs were “clear for all to see”, therefore no need for evaluations. However, these solutions mainly related to the fact that line managers and senior clinicians could now monitor users’ performance through the monitoring features of the CDSSs and could also ensure adherence with guidelines because of the compulsory nature of the CDSSs. The key stakeholders’ perceptions of the CDSS

evaluations were another barrier to evaluations. For example, T1LeadConsultant revealed that he had considered carrying out a before and after study following T1's implementation, looking at VTE audits and rates of hospital acquired VTE pre and post T1 implementation. However, this project was abandoned in part because of lack of funding but also because two other NHS Trusts, which were amongst the leaders in the VTE Exemplar Network, had published their data and T1LeadConsultant felt that the study Trust could not add "anything new to the discourse". Similarly, T2LeadConsultant and T3LeadConsultant took the evaluations that had been carried out at the developing Trusts at face value and expected to gain similar benefits as had been reported by their peers at the developing Trusts. The views of the Lead Consultants and other key stakeholders revealed a general consensus that NHS Trusts are generally similar and what works in one Trust is bound to work in other Trusts because they are structured in similar ways and follow the same guidelines.

This section has looked at a range of technical, methodological, organisational and human barriers to CDSS evaluations. The evaluators had limited experience, expertise and methods to carry out the necessary evaluations. The attitudes of the evaluators and other key stakeholders towards evaluations limited the range of and extent of evaluations that were carried out. A key contributory factor was that the three CDSSs were treated as panaceas to existing problems in the study Trust and untested assumptions were made about their benefits. There were no clear customers of the evaluations, which resulted in minimal scrutiny of both internal and external evaluations. External evaluations were taken at "face value", and their validity and fit with the study Trust were not assessed. Table 9.2 highlights some of the main barriers to CDSS evaluations that were noted in the case

studies and the enabling factors that may help to minimise the barriers and facilitate effective CDSS evaluations.

## Barriers to CDSS evaluations

- Untested assumptions that CDSS is a panacea
- Evaluations viewed as waste of time and resources
- Evaluation not seen as a priority
- Concerns about evaluation disincentives
- Evaluations seen as hindrance to CDSS adoption
- Competing priorities for resources (lack of funding)
- Concerns about raising unnecessary opposition
- Evaluation may be used as a political tool
- Concerns around questioning decision makers
- Limited availability of evaluation methods
- Lack of scrutiny and usage of previous and external evaluations (lack of evaluations of evaluations)
- Novelty of CDSSs make them challenging to evaluate
- Complexity of healthcare organisations
- Mismatch between research and practice
- Key stakeholders do not see need to evaluate
- Concerns about encouraging opposition
- Primary focus on formative evaluations
- Focus on single outcomes (pressures from internal and external evaluation customers)
- Inappropriate evaluations methods used (lack of knowledge and organisational capacity for evaluations)
- Lack of clear evaluation customers
- Evaluation customers may be “lost” due to organisational changes or staff turnover
- Interdisciplinary crossovers may result in loss of ownership of evaluations
- Poor dissemination of evaluations

## Enablers of CDSS evaluations

- Clearly identify evaluation customers, their information needs and decisions which will be affected by evaluations
- Publicise the benefits of CDSS evaluations to all key stakeholders and encourage buy in (minimise barriers)
- Show how evaluations support organisational and clinical decision-making
- Evaluate CDSSs as part of self-review processes throughout every stage of their lifecycle
- Develop a progressive organisational culture where evaluations are encouraged
- Provide adequate funding for evaluations
- Ensure that evaluators are independent and free from vested interests and hidden agendas
- Ensure that the evaluators have the expertise and capacity to carry out the required evaluations
- Publicise evaluation methods and encourage interdisciplinary contributions
- Highlight and address failures where identified and show the lessons learnt for future adopters and evaluators
- Disseminate evaluation findings to all key stakeholders
- Develop partnerships with other healthcare organisations, academic and research organisations
- Reduce the research-practice gap by encouraging the adoption of evidence-based practice
- Build organisational capacity and knowledge to select and use the most appropriate methods
- Incorporate evaluations into wider IT and business strategy and communicate the vision

Table 9.2 Barriers and enablers CDSS evaluations

The majority of the barriers to CDSS evaluations noted in Table 9.2 primarily related to the attitudes of the key stakeholders, most whom did not see the need to evaluate the CDSSs. There appeared to be an inherent resistance to carrying out any evaluations on the basis that the CDSS benefits were clear. To these stakeholders, undertaking evaluations was a waste of limited resources, which they felt could be better used elsewhere. Also, the existing professional and organisational hierarchies limited the level of scrutiny that the CDSS evaluations were subjected to. Most of the evaluators were senior physicians, who were also departmental leaders responsible for the clinical conditions that were supported by the CDSSs. It also appeared that the compulsory nature of the CDSSs and the involvement of key opinion leaders in their evaluations may have influenced the findings of these evaluations. For example, some clinical departments were reported to have complied with the CDSSs when it became apparent to them that they would be penalised operationally and financially if they did not comply. They were also reported to be reluctant to criticise or highlight the shortcomings of the CDSS evaluations because they worked closely with the evaluators and did not want to be seen to be criticising them or publicising the weaknesses of the CDSS projects and the evaluations. As such, the existing power dynamics were a key barrier in themselves. It was not clear whether the evaluators were aware of these power dynamics and their potential contribution to the general lack of scrutiny of evaluations.

The lack of openness around some of the key evaluations may be seen as a deliberate effort to keep evaluations away from some of the stakeholders. This could be viewed as inappropriate use of power to suppress any criticism and push through CDSS projects and publish only the evaluations or aspects of evaluations which produced the answers that suited the evaluators and decision makers. The reticence around evaluations may also be



viewed as showing the organisation or evaluators' reluctance to publicise any evaluations that showed negative outcomes and may result in unnecessary scrutiny and criticism or even litigation. Such concerns could be addressed by using robust methods to evaluate CDSSs at every stage of their lifecycle and disseminating results to all key stakeholders. Where failures are identified, the organisation and evaluators should highlight and address them, showing what lessons were learnt to ensure that future CDSS developers and evaluators do not repeat the same mistakes. However, there was a limited range of evaluation methods that were used, which in turn limited the evaluation portfolios of the CDSSs. It could be argued that the evaluators did not see the need to expand the range of methods because the purposes of their evaluations were limited to the immediate issues at hand, i.e., collecting enough information to report to their compliance with guidelines to regulatory authorities. However, there does appear to be a limited range of evaluation methods and guidelines within NHS settings. Furthermore, there was lack of awareness of the wider implications of the CDSSs on organisational and social issues. This may be because these issues were not considered as priorities in comparison with ensuring the technical and clinical efficacy of these CDSSs and ensuring that they adhered to clinical guidelines. Also, there was lack of interdisciplinary collaboration which could have helped to engage with and involve others stakeholders from different traditions within the study Trust and also those from partner organisations such as the affiliated local universities. The lack of evaluation documentation and poor dissemination of CDSS evaluations may have denied some of the key stakeholders within and outside the study Trust opportunities to appraise these evaluations and use them to make more informed decisions within their own settings. In particular, other NHS Trusts who may be looking to adopt or evaluate

similar technologies would stand to benefit from these evaluations if they were made available.

## **9.3 Research findings that support the CDSS evaluation framework**

In line with the literature review findings, there was consensus amongst the key stakeholders that the three CDSSs had the potential to facilitate clinical and operational improvements in the study Trust. The three CDSSs themselves were fairly novel; T1 being the first such system to be developed and implemented in an NHS hospital; T2 being the first such system to be made commercially available in the UK and T3 being the first system to be implemented across an NHS hospital. The findings that support the CDSS evaluation framework (see Figure 9.1) are now discussed in turn.

### **9.3.1 Organisational, contextual and social issues that support the literature review model**

The context of evaluation has been identified as essential to fully understand CDSSs and their effects. However the literature review revealed that most CDSS evaluation studies often ignore the context of evaluation and thus fail to fully comprehend the complex interplay between CDSS and the different actors and the organisational settings into which they are introduced. Mumford (1983) argued that despite the technical nature of computerised information systems, developers of these systems should consider human factors and aim to enhance users' job satisfaction before introducing computerised systems. Mumford argued that it was critical to assess how computerised systems would affect users in order to achieve the desired outcomes. Kaplan and Harris-Salomone (Kaplan

and Harris-Salamone, 2009a) also noted the challenges relating to the technical issues brought by health information technologies (including CDSSs), particularly problems around their functionality and interoperability with legacy systems and integration with clinical workflow. They argued that most problems were related to sociological, cultural and financial issues and thus more managerial rather than technical (Kaplan and Harris-Salamone, 2009a).

The desired outcomes were reported to have been achieved for all three CDSSs despite the main focus of the primary evaluations having been on their technical aspects. The complexity of the new CDSSs and interrelationships with organisational and social issues introduced new sociotechnical aspects that were difficult to evaluate. For all three CDSSs, there was a tendency by evaluators to look at the wider service provision in relation to guidelines and clinical pathways, while ignoring organisational and individual user needs. By primarily focusing on evaluations that looked at technical efficacy, the evaluators failed to pay attention to the users' needs, which inevitably resulted in poor 'buy in' and unintended uses of the CDSSs. It was unclear whether these shortcomings were due to the inadequacies of the evaluation methods, the narrow purposes of evaluations or both. It is likely that the purposes and motivations for evaluation certainly limited the range and boundaries of evaluation methods that were used. A modified version of Kaplan and Norton's (1995) balanced scorecard has been suggested as an alternative to cover the wider perspectives of evaluations at technological, economic, organisational and individual levels for information system projects (Land, 2000; Martinsons et al., 1999). At the CDSS level, the balanced scorecard would cover the customer, financial, internal efficiency, learning and growth opportunities and employees' perspectives. Similarly, the Health Technology Assessment (HTA) method could have helped to provide different decision-

makers with information from various angles such as the benefits, risks and costs of an intervention. However, it could be argued that the peer reviews and audits provided adequate information to the project leaders and key stakeholders and thus there was no motivation to look at the wider effects of the CDSSs.

Walsham (2006) noted that the increasingly pervasive and complex nature of information systems and their use as tools for organisational change entail new demands from evaluation activities. This was evident in the study Trust's shifting organisational priorities, which in turn led the shifts in the focus of the CDSS evaluations. The changes in the organisational priorities and subsequent changes in the range and boundaries of evaluations were primarily directed by changes in national priorities at the time. In response to these changes, the evaluators adapted the existing evaluations to suit the new regulatory reporting requirements regardless of whether there was any correlation to the CDSS and without due consideration for their effects on practice and patient outcomes. For example, the CQUIN VTE framework required all NHS Trusts to change the baseline for the achievement of minimum VTE risk assessments from 90% to 95%. The study Trust (and other NHS Trusts) responded by excluding certain patient cohorts on the basis that they were at low risk of developing a hospital acquired VTE and thus did not require a risk assessment. The study Trust immediately achieved the minimum requirement of 95% VTE risk assessment compliance and continued to secure the related CQUIN operational funding. For an outsider, it would appear that corresponding changes had been implemented to achieve these improvements. However, there had been no change at all in clinical practice and also no corresponding improvements in patient outcomes. Similarly, wider evaluations such as the Critical Care Outreach Team's service evaluation, Dr Foster and IQNARC audits all attributed substantial improvements to T3 without any evidence of

correlation or regard for simultaneous improvement initiatives that had been implemented in the study Trust. It would appear that the evaluators were under constant pressure to keep up with internal and external reporting requirements and in response, they focused on carrying out only the evaluations that would adequately provide the necessary answers and would also be considered meaningful by the decision-makers and funders, who were the key customers of the evaluations. The top-down approach to evaluations filtered from the regulatory authorities to the Trust executives, who in turn passed the responsibility to clinical leaders and their respective project teams. Little consideration appears to have been taken of the effect of these pressures on other stakeholders, especially the users of the CDSSs. It would appear that the lower the stakeholders were in the decision-making ladder, the lower the ownership with respect to CDSS evaluations. Also, the lower the evaluators were in the decision-making hierarchy, the lower the impact of their evaluations on decision-making. For example, the VTE Link Nurse-led audit of the study Trust's adherence with the NICE VTE Quality Standards found huge disparities between the reported and actual compliance. It would appear that the method that was being used to collect the monthly VTE CQUIN audit data primarily focused on the quantity rather than quality and validity of risk assessments. Despite these findings, there were no changes in the way monthly VTE risk assessment data were collected and the study Trust continued to report that they were achieving over 95% valid VTE risk assessments. It could be argued that this evaluation failed to make an impact on decision-making because it was not commissioned by the key decision-makers or that its findings were in contrast with the study Trust's priorities at the time. It was also clear that VTELeadNurse and the VTE Link Nurse Network lacked capacity and authority to effectively implement changes in response to findings from evaluations.

Most clinical leaders and senior managers generally agreed that the CDSSs and computerised information systems in general had brought radical changes that had impacted positively on clinical processes and patient outcomes. There were positive perceptions about how these technologies had enabled the modernisation of the clinical services which they supported and provided support for the clinical teams to cope with the increasing demand for healthcare services. Senior managers and clinicians expected long term benefits such as improved patient safety and care, increased efficiencies, more effective and appropriately delivered healthcare in line with the recommendations by the Institute of Medicine report (Kohn et al., 2000b) and subsequent patient safety studies (Leape and Berwick, 2005; Stelfox et al., 2006). However, some nurse leaders expressed concern about the potential deskilling of nurses, the negative effects of overreliance on CDSSs and over-allocation of limited funding to these technologies at the expense of adequate training and ensuring that the skill mix on the wards was maintained. Other key stakeholders felt that all the CDSSs were inappropriately specified and unsuitable for some clinical areas. Even a small scale system like T2 was deemed to be unsuitable by some of the users. T3 lacked acceptance and trust by users in part because of poor specifications during the first few years of implementation which resulted in disruptions in clinical workflow, job dissatisfaction and potential harms to patients. Similarly T1 was viewed as an unnecessary and bureaucratic process which did not benefit patients or the users. Indeed, all three CDSSs were not fully integrated with existing legacy systems despite the original design expectations. However, none of the evaluations were geared to pick up the concerns of these stakeholders and any concerns raised were generally considered by the project leaders to be part of the expected resistance to any new innovations. Berg et al. (1999) noted the sociotechnical sense of computerised health information technologies,

whereby success of any system is measured on the basis of technological artefacts and the actors who use the technology to make up the Information System. It would appear that none of the evaluations formally assessed the human-CDSS interaction and thus missed opportunities to understand the wider effects of the CDSSs beyond the technical aspects and stated objectives.

### **9.3.2 Purposes of evaluations that support the literature review framework**

The literature review showed that CDSS evaluation purposes varied considerably. The main purposes of CDSS evaluations identified in the literature review include technical and clinical efficacy; user performance and accuracy of clinical decision making; satisfying regulatory authorities; assessing cost effectiveness and justification of IT investment to key stakeholders. It has been argued that the purposes of CDSS evaluations should include a range of technical, clinical, organisational, human and social issues (Kaplan, 2001a; Kaplan, 2001b).

The main purposes of the CDSS evaluations in the three cases were to gather information that would satisfy regulatory reporting requirements for the clinical conditions and services that were supported by the CDSSs. Other evaluations were also carried out to assess the technical and clinical effectiveness as well as to monitor the performance of the users. The evaluators were primarily concerned with the CDSSs' immediate usability needs, especially proving that the CDSSs worked as intended, were compliant with guideline recommendations, produced reliable information and recommendations and ensured the security of patient data. The senior clinicians were keen to ensure that the software algorithms correctly translated the underlying clinical guidelines, which was assessed through verification and validation of the CDSSs and their underlying algorithms. These

evaluations are essential at the developmental and early implementation stages to show technical and clinical efficacy. However, as shown in the case studies, proving technical and clinical efficacy does not necessarily result in the CDSSs being used as intended. These evaluations were in line with the literature review findings (Kaplan, 2001b; Kaplan and Maxwell, 2005; Garg et al., 2005b; Bright et al., 2012). The majority of the evaluations did not consider the complexities of introducing the CDSSs into the study Trust, especially the effects of the attitudes of the key stakeholders and the human-CDSS interaction. The attitudes of the evaluators about CDSS evaluations also affected the evaluations that were carried out, although most of the evaluators considered their evaluations to be “objective” and based on the positivist tradition. All three CDSSs were evaluated by the same people who had developed them, which brought questions around their objectivity. Also, these evaluators were senior clinicians who had reporting responsibilities to the study Trust Board and regulatory authorities. These CDSSs had been commissioned with set objectives to generate information that would be sufficient to report to the regulators and commissioners. As such, their evaluation focus was driven by the need to achieve the immediate intended goals rather than assessing the wider effects of the CDSSs.

### **9.3.3 Approaches and methods for CDSS evaluations that support the literature review model**

The literature review showed that CDSS evaluation approaches were broadly classified under positivist, interpretivist and critical assumptions (Greenhalgh and Russell, 2010b). Also, most CDSS evaluation studies are positivist, and primarily use the RCT method as the “gold standard” on the basis that it is bias free and cause and effect can be objectively and precisely measured (Greenhalgh and Russell, 2010b; Friedman and Wyatt, 2006; Kaplan,



2001a). However, the majority of systematic reviews found that RCT-based CDSS evaluation studies were widely heterogeneous and challenging to compare and generalise (Garg et al., 2005b; Randell et al., 2007; Bright et al., 2012). Consequently, most systematic reviews resorted to narrative synthesis rather than pooling results or meta-analysis, which often resulted in inconclusive results and cautious recommendations for practice (Kaplan, 2001b). Interpretivist methods have also been used to evaluate CDSSs (Greenhalgh et al., 2009). These methods object to the contention that reality is objective and can be precisely measured. Instead, they look at reality as a social construction which is enriched by the researcher's identity and values throughout the research process. Critical approaches seek to gain an understanding of the power relations within organisations that are manifested through the interests of the different stakeholders (Klecun and Cornford, 2005). This understanding is achieved through systematic questioning of the status quo. Although interpretivist and critical approaches to CDSS evaluations are gaining popularity, they are still not widely used in comparison with positivist approaches.

The literature review highlighted that quantitative methods did not explain context-specific issues regarding fit of CDSS, whether CDSSs are used or not and their acceptance among other key attributes. Although RCTs are generally reported to be the "gold standard" for CDSS evaluations and their popularity in clinical environments, none of the three CDSSs were evaluated using RCTs at any stage of their lifecycle. Indeed, none of the project leaders even considered using RCTs. This may be because they are expensive to set up and challenging to manage, and the expertise and other resources were not available. More precisely for the three CDSSs, the project leaders and other key stakeholders had resolved that the CDSSs would be implemented anyway regardless of whether or not the evaluations had been carried out using the "gold standard". Instead, the majority of the evaluations

were based on peer reviews and simple before and after audits which were fraught with limitations and confounding factors that could have been controlled using RCTs or similar methods. The literature review also showed that most CDSS evaluations primarily carried out summative evaluations, particularly those that related to the implementation/installation of the CDSS projects. However, this was not the case for the three CDSSs. There were no longitudinal evaluations over time apart from the second benefits realisation study for T3. For all three CDSSs, the majority of evaluations were formative and focused on single issues during development and early implementation stages. For all three CDSSs, there was minimal effort to engage with the key stakeholders (especially users) to understand their information needs. Indeed, none of the evaluations adopted user satisfaction as an evaluative measure at any phase of the CDSSs' lifecycle.

There was also little evidence of organisational readiness to effectively assess the CDSSs both in terms of organisational capacity to carry out the evaluations and the expertise to carry them out effectively. Formal assessments of the existing resources, in the form of legacy systems infrastructure and human resources to carry out the evaluations before the CDSSs' implementation would have revealed the gaps in evaluation expertise. It would have been likely that had this been raised as an issue at this early stage, the Trust Board and Commissioners would have been more prepared to allocate adequate funding for evaluations, even if it meant externally sourcing the required expertise because they were keen for the CDSS projects to be successfully implemented. Despite the availability of previous audits of the "old" paper-based systems and legacy systems, none of the CDSS evaluations used them as comparators to formally evaluate the effects of the CDSSs. Previous evaluations would have helped to inform current evaluations as well as providing a baseline to measure the effects of the CDSSs, while also separating some of the

confounding factors. However, this may have been made challenging by the lack of interaction between the CDSSs and existing legacy systems and in some cases, continued use of non-computerised systems in parallel with the CDSSs. There was also no effort to assess the fit of evaluation methods to the respective CDSSs and the study Trust. Even where explicit blind spots were noted in the evaluation methods, it would appear that there were no efforts to address them as long as the limitations did not affect the ultimate goals of collecting the audit data for reporting to regulatory authorities. Also the failure to undertake any formal summative evaluations was a lost opportunity to assess the wider effects of the CDSSs at every stage of their adoption lifecycle. This was in contrast to the literature review, which suggested that most CDSS evaluations were summative rather than formative. There was also lack of scrutiny of external evaluations even where they were critical to inform adopting decisions and large scale implementation. In the case of T3, this led to a CDSS that was “unsuitable” for the study Trust being adopted and resulting in additional costs and major delays in large-scale implementation. Although the study Trust reported that the expected benefits (as reported by the developers) had been achieved, there was lack of evidence to support these claims and these evaluations had not been published or disseminated to most of the key stakeholders. This may be because there was inadequate evidence to support these claims or that the evaluators did not wish to share this information with the other stakeholders. There was therefore lack of information about the wider effects of the CDSSs on the reported outcomes and the wider organisation. These findings concurred with the literature review findings, which showed that CDSS evaluations may sometimes be used to pursue the interests of the evaluators and key decision makers (Rigby, 2001; Friedman and Wyatt, 2006). For the three case studies, these

interests primarily related to adopting evaluation methods that were likely to assess the CDSSs' adherence with NICE guidelines and report such findings accordingly.

### **9.3.4 Benefits of evaluations that support the literature review model**

The literature review found ethical, economic, professional and organisational benefits of evaluating CDSSs (Rigby, 2001; Rigby et al., 2001; Friedman and Wyatt, 2006; Liu and Wyatt, 2011a). Most studies have generally reported benefits of CDSSs for user and system performance, but few studies have looked at patient outcomes, and even fewer have reported improvements in patient outcomes (Jaspers et al., 2011b; Buntin et al., 2011; Black et al., 2011; Hunt et al., 1998; Lau et al., 2010). However, few CDSSs evaluations looked at patient outcomes and even fewer reported benefits for patients. The novelty of CDSSs has been reported to bring about additional risks for users and patients. Healthcare professionals and organisations therefore have a moral duty to evaluate them effectively to assess that they provide the intended benefits. Additionally, these benefits need to be shared with all the key stakeholders, including the users and the wider research community. Furthermore, CDSSs are expensive to develop and procure. It has been noted that evaluators and decision makers often carry out evaluations that retrospectively find benefits to justify investing in these projects.

The main benefits of CDSS evaluations were the ease of access to information that had previously been difficult to obtain using paper-based systems. Obtaining this information was important for the respective project leaders and clinical managers because they could now monitor CDSS users' clinical performance and implement remedial changes where indicated. It also became cheaper and quicker to use CDSSs to make retrospective analyses for processes such as the root cause analyses and routine audits. Evaluations such as the

VTE CQUIN audits had a dual purpose of reporting the study Trust's performance to the commissioners and to identify departments that were non-compliant with the NICE guidelines. The T1 project team worked with some of the departments that were failing to meet the minimum standards and remedial action plans were put in place to make improvements. However, the downside was that the T1 project team was under-resourced to effectively address all the issues that were being brought up by these audits. For example the audit that was carried out by VTELeadNurse and the VTE link nurses revealed various shortcomings in all the clinical areas that were audited. However, the evaluators had very limited resources and could only follow up the cases where patient safety had been put at risk. Although there were clear benefits for carrying out these audits, it became a futile exercise because the evaluators did not have adequate resources to develop and implement remedial actions. Also the decision makers in the affected departments had other competing priorities at the time which required their attention and the necessary resource allocation.

### **9.3.5 Barriers to CDSS evaluations that support the literature review model**

The literature review highlighted human, organisational, technological and methodological barriers to CDSS evaluations. Boland et al. (2014) noted that despite the proven benefits of evaluating all innovations (including CDSSs), there were significant barriers such as lack of a gold standard for evaluations and scarcity of knowledge regarding user needs prior to health information technology development. They also noted the existence of complex sociotechnical aspects relating to these systems and increasingly multi-stakeholder teams that could affect the achievement of intended outcomes of health information technology interventions.

The CDSSs' effects on existing communication channels, clinical workflow and interactions between stakeholders, potential effects of shifts in responsibility, and different value sets between different stakeholders were noted in the literature review as the key factors that were likely to impact on the effectiveness of CDSS evaluations. However, these effects were not evaluated in all three cases. There was also lack of CDSS project integration with the existing work environments and the wider effects of regulatory and other policy requirements. Such integration was reported by some of the project leaders to be technically challenging, especially because of the heterogeneity of the various clinical departments and disparate attitudes of the key stakeholders. Others argued that CDSS integration with existing legacy systems was not difficult to implement but instead, it was hindered by the key stakeholders' concerns around the costs and information governance issues. There was also minimal evidence of alignment of the CDSSs with the study Trust's strategic and business goals, and their lack of integration with routine clinical practice and clinical workflow. Instead, there was evidence of CDSSs being viewed by users as top-down obstacles that had to be worked around, and not as part of routine clinical practice as intended by the CDSS adopters. It appeared that once the departmental clinical leaders and their teams realised that the CDSSs were compulsory, they devised ways to minimally adhere to them without compromising their diverse and often competing interests. Little attention appeared to have been paid to the information needs and legitimate interests of some of the stakeholders. Indeed, the project leaders acknowledged that some clinical departments did not require the CDSSs but they had to find ways of implementing them in these areas because census reporting was required for all departments. In such cases, it could be argued that the departmental leaders' concerns that the CDSSs were administrative tools for data collection and management reporting rather systems to

improve clinical processes and patient outcomes were justified. It would appear that the CDSSs and their related evaluations were being used for legitimising the vested interests of some of the key stakeholders and as tools to sway decision-making in favour of continued funding for the CDSS projects. The lack of formal consultation with the key stakeholders prior to adoption of the CDSSs, especially the users, is likely to have caused the poor 'buy in' and lack of CDSS ownership. Although T1 project leaders made efforts to engage with all the relevant departmental leaders prior to its implementation, the opinions of these departments did not really matter because the adoption decision had already been made and the consultation process was a mere formality to agree terms on which T1 would be implemented. The compulsory nature of all the three CDSSs and the top down approach to their adoption alienated key stakeholders who inevitably viewed the CDSSs as part of "government diktat" and management control over their work, with little or no effect on clinical practice and patient outcomes. As noted in the literature (Boland et al., 2014; Berg, 1999; Kaplan, 2001b), introducing CDSSs requires a good understanding of user needs and the complex social aspects and interrelationships across multi-disciplinary stakeholders. There was lack of consideration for these complexities in the evaluations that were carried out. Indeed, none of the formal evaluations even alluded to the existence of these complexities, although the project leaders and other key stakeholders acknowledged them during interviews. CDSS users did not share the ultimate vision of the project leaders regarding the CDSSs and felt that these systems had been imposed on them without regard to their effects on their clinical practice and workflow, user performance and patient outcomes. This was especially so for the specialised clinical departments, some of which already had more complex computerised health information systems and did not see the need for these CDSSs. It was therefore no surprise that the majority of CDSS users were

“just doing enough to tick the box and cover their backs” because they did not regard them as useful in their clinical areas. Instead of helping to make clinical decisions and achieve the intended clinical outcomes, the CDSSs thus became administrative chores that had to be worked around to cover oneself against potential penalisation or litigation.

The CDSSs’ novelty and the widely held views that they were panaceas to existing clinical, operational and financial problems made them challenging to evaluate. Indeed, some key stakeholders did not even see the need to evaluate the CDSSs because they perceived that their introduction as replacements for the previous paper-based systems had led to “overwhelming” improvements. However, these improvements were primarily related to the availability of previously inaccessible audit information regarding user performance rather than the benefits of the CDSSs themselves. It would also appear that evaluation methods were selected for their ability to generate information that would be reportable to regulatory authorities and to gain financial incentives from Commissioners. There did not appear to be any effort to learn from previous CDSS evaluations within or outside the study Trust or to consult wider CDSS or computerised health information technology evaluation literature. Essentially the CDSSs were being assessed for their worth as management information systems, change agents and reporting tools to senior management and regulatory authorities rather than their effects on clinical practice, user performance or even patient outcomes. Systematic reviews by Garg et al. (2005b) and Kawamoto et al. (2005) showed that CDSSs that were introduced as change agents were unlikely to be successful. Other barriers related to the CDSSs’ novelty included the lack of previous CDSS evaluation experience for the evaluators and the complexity of carrying out CDSS evaluations. The evaluators were not well informed about how best to evaluate CDSSs and there was no effort to assess or improve organisational readiness for CDSS evaluations.



Other human barriers to CDSS evaluations included the lack of scrutiny of peer evaluations. There appeared to be a tacit understanding amongst the key stakeholders that the CDSSs had been evaluated appropriately even where there was no evidence to support these assumptions. The majority of the evaluations had limited documentation and in some cases, there was no evaluation documentation at all. The interviews revealed that CDSS project leaders' personal and professional judgements significantly influenced and at times replaced formal evaluations. This was particularly so with regards to the successes of the evaluations that had been carried out and the effectiveness of the CDSSs. However, without any assessment of cause and effect, it was impossible to make a distinction between the CDSSs' effects and those of other improvement programmes which had been implemented in the study Trust at the same time. Furthermore, all the primary evaluations had been carried out by the project leaders themselves, thus raising questions regarding their objectivity because they were involved with every aspect of the adoption/development, implementation and evaluation processes. There were no structures in place to carry out evaluations of the evaluations that had been carried out or question the validity of the reported outcomes. There was also no scrutiny of external evaluations that had been carried out by the peers of the three CDSS project leaders. It would appear that the opinions of the peer groups and leading figures within the physicians' collegiate groups had a significant effect on the range of evaluations that were carried out. For example, T1LeadConsultant discontinued a planned evaluation looking at the impact of T1 implementation because two leading NHS Trusts had published similar studies and he did not feel that the study Trust would "add anything new to the discourse". Furthermore, T1LeadConsultant felt that the study Trust was likely to have achieved similar results to the publishing Trusts. T2LeadConsultant and T3LeadConsultant accepted the evaluations that

had been carried out by their peers at the respective developing NHS Trusts because in part, they trusted their peers and they were likely to carry out similar evaluations themselves. All three Lead Consultants and other key stakeholders also disagreed with some elements of the NICE guidelines and other Department of Health initiatives relating to their respective clinical specialties (which informed the CDSSs) but did not formally raise any concerns. There appeared to be a tendency to work with, rather than work against the recommendations from collegiate and regulatory authorities.

## 9.4 Key research findings

This research project sought to answer the following questions in the context of a typical NHS hospital:

- 1 What are the key factors that affect CDSS evaluations in a typical NHS hospital setting?
- 2 How do these factors relate to the CDSS evaluation framework that was developed from the CDSS literature review?
- 3 To what extent do evaluations affect decisions to adopt CDSSs in healthcare settings and which evaluation methods are most likely to inform CDSS adoption decisions and why?

Table 9.3 shows a summary of the key findings from this research. Each of these findings will now be discussed in turn.

Category of findings	Summary of key findings
<i>Contextual factors</i>	<ul style="list-style-type: none"> <li>• There was limited consideration of the context of CDSS adoption and evaluation, with assumptions and generalisations about CDSSs</li> <li>• Lack of financial awareness amongst nurses, e.g., costs related to CDSSs and economic benefits</li> <li>• There were concerns around the potential for CDSSs resulting in the deskilling of staff, especially nurses</li> <li>• There was limited consideration about how CDSSs could potentially affect workflow and the quality of the work environment</li> </ul>
<i>Understanding of evaluations</i>	<ul style="list-style-type: none"> <li>• In concurrence with the literature review, there was limited understanding of evaluations, i.e., what evaluation is; who evaluates; how to evaluate, when to evaluate and how long to evaluate for</li> <li>• Ongoing service evaluations were presented as CDSS evaluations, even where there were no direct links between the specific aspects of the services and CDSSs</li> <li>• There was a belief that successful CDSS implementation was a solution to existing problems and equated to evaluation, even when such benefits had not been proven, there was an expectation that they would be achieved in the future</li> <li>• Improved data visibility was assumed to be actual improvements in workflow and patient outcomes without any evidence</li> </ul>
<i>Methodological factors</i>	<ul style="list-style-type: none"> <li>• There were limited evaluation methods used, with audits being the method of choice</li> <li>• There was limited understanding of the CDSSs and methodological issues around their novelty</li> <li>• Due to CDSS novelty, there were limited reference points for evaluators and little or no reference to CDSS evaluation literature</li> <li>• Evaluators focussed on immediate regulatory compliance requirements and selected evaluation methods that would best generate the required information rather than appropriateness of methods</li> <li>• Formal evaluations did not look at how CDSSs would affect existing work patterns, shifts in responsibilities and interoperability with existing legacy systems</li> <li>• Evaluations had explicit 'blind spots' that emanated from decision makers' primary focus on issues which they considered important at the time. Although in some cases there was awareness of such 'blind spots', remedial action was not taken</li> </ul>

	<ul style="list-style-type: none"> <li>• Claims that were made about economic benefits were based on anecdotal evidence, opinions and professional judgment in the course of key actors' routine work rather than real evidence of CDSS success</li> <li>• Choice of evaluation method (s) was influenced by peer group</li> <li>• Assumptions about CDSS benefits were not backed by evidence, 'after this, therefore because of this' syndrome</li> <li>• Despite the widely held belief that RCTs and experimental methods were the 'best way' to evaluate CDSS, these methods were not used or even considered in any of the evaluations investigated</li> <li>• In contrast to findings from the literature review, most CDSS evaluations investigated were formative rather than summative</li> <li>• There were no direct comparisons between CDSSs and existing legacy systems or paper-based alternatives, even where the data were readily available</li> </ul>
<i>Technical factors</i>	<ul style="list-style-type: none"> <li>• Little or no attention paid to CDSS modularity and user friendliness in the evaluations</li> </ul>
<i>Evaluation purposes</i>	<ul style="list-style-type: none"> <li>• Evaluation priorities shifted throughout CDSS lifecycle</li> <li>• Evaluations were primarily focused on guideline adherence and protection of organisational reputation</li> </ul>
<i>Benefits of evaluations</i>	<ul style="list-style-type: none"> <li>• There was limited evidence of CDSS benefits</li> <li>• There were frustrations due to unmet expectations</li> </ul>
<i>Human factors</i>	<ul style="list-style-type: none"> <li>• Some stakeholders noted power issues, particularly inappropriate use of power to gain compliance</li> <li>• Evidence of evaluations being used to promote vested interests</li> <li>• There was a difference between how and how much CDSSs were being used, and how they were actually being used in practice</li> <li>• Resistance to CDSSs, and the change elements they were associated with were not investigated</li> </ul>
<i>Organisational factors</i>	<ul style="list-style-type: none"> <li>• There was lack of commitment to evaluations, manifested by the lack of funding for evaluations, selective use of evaluations and sometimes ignoring evaluations completely</li> </ul>

	<ul style="list-style-type: none"> <li>• CDSSs were used as change agents with little regard to their effects on staff morale, workflow and other wider organisational issues</li> <li>• There was widespread lack of project and CDSS documentation</li> <li>• Although there was awareness of CDSS underutilisation, there was no action taken to improve or encourage their usage</li> <li>• Key stakeholders highlighted issues around multiple initiatives</li> <li>• Concerns raised about top-down approach to CDSS adoption and evaluations and the related pressures</li> <li>• There was lack of CDSS and evaluation ownership</li> <li>• Evaluation customers were not always clearly identifiable, and at times organisational changes resulted in complete ‘disappearance’ of the evaluation customers</li> <li>• The lower the stakeholders were in the organisational hierarchy and decision making ladder, the lower the ownership with respect to CDSS evaluations. Also the lower the evaluators were in the decision-making hierarchy, the lower the impact of their evaluations on decision making</li> <li>• There was limited organisational readiness for CDSS adoption and evaluation (lack of robust infrastructure)</li> </ul>
<i>Barriers to evaluations</i>	<ul style="list-style-type: none"> <li>• Project leaders reported the lack of funding to effectively carry out ‘meaningful’ evaluations</li> <li>• Project leaders noted that it was easy to ‘cheat’ evaluations, whereby CDSS users at times did the bare minimum to ‘tick the box’ and be seen to be following the related guidelines</li> <li>• Some key stakeholders reported facing ethical dilemmas, for example, they had limited resources to effectively follow up evaluation findings and implement remedial actions in clinical areas that were failing to meet the guideline requirements</li> </ul>
<i>Use of evaluations</i>	<ul style="list-style-type: none"> <li>• Information from evaluations primarily used to satisfy external stakeholders, especially regulatory authorities and commissioners</li> <li>• There was limited use of information from evaluations to improve internal processes and patient outcomes</li> <li>• Some key stakeholders were concerned that other NHS Trusts were employing ‘clever reporting’ techniques to be seen to be meeting the national targets, even where there were no corresponding improvements in patient outcomes</li> <li>• There was little or no dissemination of CDSS evaluations to key stakeholders</li> </ul>

	<ul style="list-style-type: none"> <li>• There did not appear to be much use of evaluation information for decision making within the study Trust</li> </ul>
<i>External influences</i>	<ul style="list-style-type: none"> <li>• Key stakeholders were concerned that although the regulatory authorities and by extension, the government set guidelines and targets, there was little evidence of enforcement of their expectations nor scrutiny of self-reported results.</li> <li>• ‘Sold benefits’ of CDSSs were not evaluated systematically. Key decision makers trusted their colleagues within their peer groups and did not question the validity of ‘sold benefits’</li> <li>• External evaluation customers (commissioners and regulators) were considered more important</li> <li>• The continually shifting national guidelines and current priorities were often in competition with organisational needs</li> </ul>
<i>Stakeholder perceptions</i>	<ul style="list-style-type: none"> <li>• The three CDSSs studied were viewed as management control systems rather than CDSSs</li> <li>• There was a general perception that CDSS evaluations were primarily done for funding and regulatory purposes, i.e., ‘what gets measured gets done’</li> <li>• There was a general perception that CDSS evaluations were not geared to provide useful information for clinical decision making</li> </ul>
<i>Key issues missed by evaluations</i>	<ul style="list-style-type: none"> <li>• There were no evaluations of unwanted, unanticipated and actual or potentially harmful effects of CDSSS following implementation</li> <li>• There was an assumption that once CDSSs had been proven to work, then they would continue to do just that without need for further evaluations</li> <li>• There was little evidence of scrutiny of external evaluations</li> </ul>

Table 9.3 Summary of key research findings

This research project carried out assessments of CDSS evaluations in real NHS clinical settings and involved the key stakeholders who were involved in adoption of the technologies and their evaluations. The evaluators largely took a positivist stance towards the evaluations, with a primary focus on the achievement of each CDSS project's stated objectives. Despite the widely held view that the RCT is the gold standard for CDSS evaluations, none of the evaluations in the three cases used this method. Most evaluations were informal, and based on the opinions of the key stakeholders who were also the project leaders.

The main purposes of the evaluations were to assess CDSS compliance with NICE guidelines and provide the means to report such compliance to the regulatory authorities. Some of the evaluations were also motivated by financial incentives from service commissioners.

There was no separation between the CDSS projects themselves and their respective evaluations. CDSS evaluations were deemed to have been successful on the basis of their successful implementation and ability to generate reportable information to the regulatory authorities and commissioners. However, there were no clear definitions of CDSS failure or failure of CDSS evaluation, in concurrence with the literature review. The success of CDSS projects and CDSS evaluations was done using simple metric-specific reports that aimed to show adherence with NICE guidelines and service commissioners requirements. However, this approach ignored the complexity of the processes and environments where the CDSSs were implemented and resulted in disincentives and failure to achieve the intended goals.

Another important finding of this research was that the attitudes of the key stakeholders towards CDSS evaluations were among the main barriers to those evaluations. This research also showed that these barriers were closely related to the existing organisational

culture, particularly attitudes towards computerised health information technologies in general and the hierarchical nature of NHS organisations. In such organisational settings, it is difficult to challenge the status quo, and evaluations may be used to protect vested interests, encourage compliance with initiatives, and as a tool to gain formal support from decision makers.

The novelty of CDSSs make them challenging to evaluate, in part because there are limited methods available, but also because evaluators may not have the necessary experience and expertise to select the most appropriate methods and use them effectively. Evaluators may inadvertently use inappropriate methods and report positive findings, even where no causal links may exist. This may result in decision makers using incorrect information to make key decisions and continue to fund or invest in similar technologies. Interdisciplinary work should also be encouraged to expand the range of evaluation methods used at different stages of the CDSSs lifecycle.

There is need for healthcare organisations to invest in CDSS evaluations, both as part of the CDSS projects and to build capacity and readiness to adopt and use CDSSs effectively, and to share their knowledge with the wider healthcare and research community. To help to build this capacity, NHS organisations need to foster relationships with other organisations that have a strong culture of CDSS evaluations. This could help to drive out the current practice of CDSS evaluations being delegated to the same people who have developed them. More independent evaluations would provide greater objectivity and counter vested interests which may cloud evaluators' judgments about what needs to be evaluated, when and what methods should be used. CDSS evaluation findings should be published, even where they produce unwanted, unanticipated or equivocal results. It is essential to discuss the results openly and show how the organisation has learnt from any errors and mistakes to



ensure that such mistakes do not recur in future and that other healthcare organisations do not make such mistakes.

There was little evidence that evaluations had any significant effect on the CDSS adoption decisions. This was true of T2 and T3 where prior evaluations had been carried out at their respective developing Trusts. Although the T2 and T3 project leaders noted that these prior evaluations were important for their decision-making, it did appear that looking at these evaluations had been a mere formality; the decisions to adopt had already been reached because both Lead Consultants trusted their peers who had developed these CDSSs and carried out their subsequent evaluations. It also seemed that the Lead Consultants were more likely to accept positivist evaluations because they are more commonly used within the medical discipline and that they would also use the same methods for their own evaluations. There was little evidence of considering alternative methods. This research showed that CDSS adoption processes were inextricably linked with evaluation. Successful implementation was reported as successful evaluation.

## **9.5 Evaluation factors that extend the CDSS evaluation framework**

Having looked at the extent to which the research findings supported the CDSS evaluation framework developed from the literature, this section looks at additional factors from the case studies that help to extend the framework. Despite the RCT being reported to be the “gold standard” for CDSS evaluations in the literature review, none of the CDSSs evaluations that were investigated used this method. The majority of evaluations were informal, and were primarily based on the opinions of the project team leaders. The

evaluators used simple before and after comparisons looking at the wider services and audits and peer reviews as their key methods of evaluation. Also, despite evidence-based medicine being a major driver for NHS interventions, reinforced through NICE guidelines and other Department of Health initiatives, the evaluations for the three CDSSs were messy, sometimes chaotic and primarily based on incidental and anecdotal evidence. Some of the key stakeholders did not see the need to evaluate the CDSSs at all because they believed that their successful implementation to replace the paper-based systems was enough. There was also lack of scrutiny of peer evaluations carried out within and outside the study Trust. There was a tacit understanding between the evaluators and the key stakeholders that the CDSSs had been evaluated appropriately even where evaluation documentation was limited or unavailable. There was also no ownership of most of the evaluations and no demands from evaluation customers for evidence of quality or robustness of evaluations. There was also widespread lack of understanding of the purposes of the CDSSs and how they should be evaluated. The expected users of the CDSSs suspected that the primary aims of the CDSSs and their related evaluations were to fulfil regulatory requirements following top down pressure from the government and regulatory authorities to Trust Executives, who then passed the responsibility to Clinical Directors to come up with solutions that proved compliance. The compulsory nature of the CDSSs disregarded the need to link these CDSSs as solutions to real clinical problems and with potential benefits to the users and ultimately patient outcomes. It appeared that the evaluators and key decision-makers were more interested in CDSSs that would provide management information for reporting to regulatory authorities than the wider effects of the CDSSs on organisational or patient outcomes. In response, there was a prevalent culture across the study Trust of CDSS users just using them enough to “cover their backs”

and “box ticking” because they did not believe that the CDSSs benefited their clinical practice or patient outcomes. The range of multi-professional stakeholders and the multiple service improvement initiatives added to the complexity of CDSS evaluations, which made it difficult to prove whether many of the reported improvements were resulting from the CDSSs or concurrent improvement efforts.

The project leaders reported that adequate funding had been allocated for the adoption of their respective CDSSs. However, it was unclear how much of this funding had been allocated to CDSS evaluations. It appeared that funding was only allocated for evaluations during the early stages of development and implementation to prove the CDSSs’ technical and clinical efficacy. Once the immediate objectives had been achieved, it would appear that they saw no benefits to be gained from funding any additional evaluations. However, it is possible that the real effects of CDSSs may not be immediately apparent and therefore other evaluations at different phases of the CDSS’s lifecycle would be needed to assess continued effectiveness. The literature review highlighted the increasing investment in computerised health information technologies to support healthcare delivery in most developed countries. Some of the expected benefits of introducing these systems include improved healthcare quality, opportunities to scale up healthcare delivery systems and to deliver financial savings. These expectations often result in increased pressure to adopt health information systems (including CDSSs) in order to achieve the expected benefits but at present there is insufficient scrutiny as to whether the benefits are actually achieved.

NHS Trusts need to understand the complexities and challenges of adopting these systems, regardless of whether they are used in a single department or across the whole Trust. Although these complexities and challenges are well documented in the literature (Garg et al., 2005b; Kawamoto et al., 2005; Kaplan and Harris-Salamone, 2009a), they did not seem

to influence the evaluation of the three CDSSs. Even a small-scale system like T2, which was only implemented in a single department was fraught with elementary failures that could have easily been addressed. Failures in CDSS evaluations, and the CDSSs themselves was reported as an under-addressed problem (Kaplan and Harris-Salamone, 2009a; Karsh et al., 2010). Indeed, none of the project leaders acknowledged any failures in the evaluations that had been carried out. For all the three cases, the definition for CDSS evaluation success was the successful implementation of the CDSS and ability to produce the required information to show adherence with regulatory requirements. This narrow definition of success limited the evaluations that were carried out and the level of scrutiny of evaluations and inevitably resulted in missed opportunities to widen evaluation boundaries and better inform decision makers.

This research has shown that the evaluations were not geared to provide adequate information to the key stakeholders and decision-makers about the CDSSs and their immediate and wider effects. Furthermore, there was little effort to engage with the intended users to establish their needs or understand the effects of the CDSSs on their workflow and the wider organisation. Also, the existence of multiple improvement efforts introduced initiative-fatigue, where nurses and physicians felt overwhelmed by the volume and demands of these initiatives. This inevitably resulted in limited connection by some of the stakeholders between the national initiatives and clinical goals in their respective departments. The evaluators could have done more to separate the effects of the CDSSs from those of the concurrent initiatives and to highlight areas of success by disseminating CDSS findings and making them relevant to the different stakeholders. Boland et al.'s (2014) mixed methods evaluation framework showed the value of integrating the assessment of user needs and user-perceived usefulness of novel technologies. Such an

evaluation framework would be appropriate for CDSS evaluations because it would bridge the gap between constantly evolving user needs and adaptiveness of technology designs during the iterative early stages of a system's development and throughout its lifecycle. This also allows CDSS users to fully review and articulate their needs and the level of decision support that is required. There is need to clearly define successes and failures of CDSS evaluations and also to take advantage of the potential opportunities to enhance the successes and to learn from evaluation failures.

## **9.6 Key contributions to CDSS evaluation literature**

This research shows that CDSS evaluations investigated took a predominantly narrow view and indicates the need for a more systemic approach. To illustrate this, the CDSS evaluation framework developed from the literature review will be synthesised with the key findings from this research. This will help to further extend the CDSS evaluation framework across all five factors identified in the literature review and emergent findings from the actual research. Figure 9.6 shows a revised version of the CDSS evaluation framework (see Figure 9.1) that was developed from the literature review. It incorporates the key contributions of this research to CDSS evaluation literature and draws heavily on Tables 1 and 2 and Figures 9.2, 9.3, 9.4 and 9.5. The CDSS evaluation framework from the literature review highlighted that the context, purposes, approaches and methods, benefits and barriers were the key factors that affected CDSS evaluations. These factors were found to be relevant throughout the research across all the case studies to varying extents. However, this research found additional factors that contribute to the synthesised CDSS evaluation framework as follows:

- Organisational strategy and vision

- Goals of the CDSS project and the decisions that are likely to be affected by evaluations
- Attitudes of the key stakeholders towards CDSS evaluations
- Organisational readiness and capacity to carry out evaluations

These factors will now be discussed in turn, including their relationship with the factors that were highlighted in the original CDSS evaluation framework.

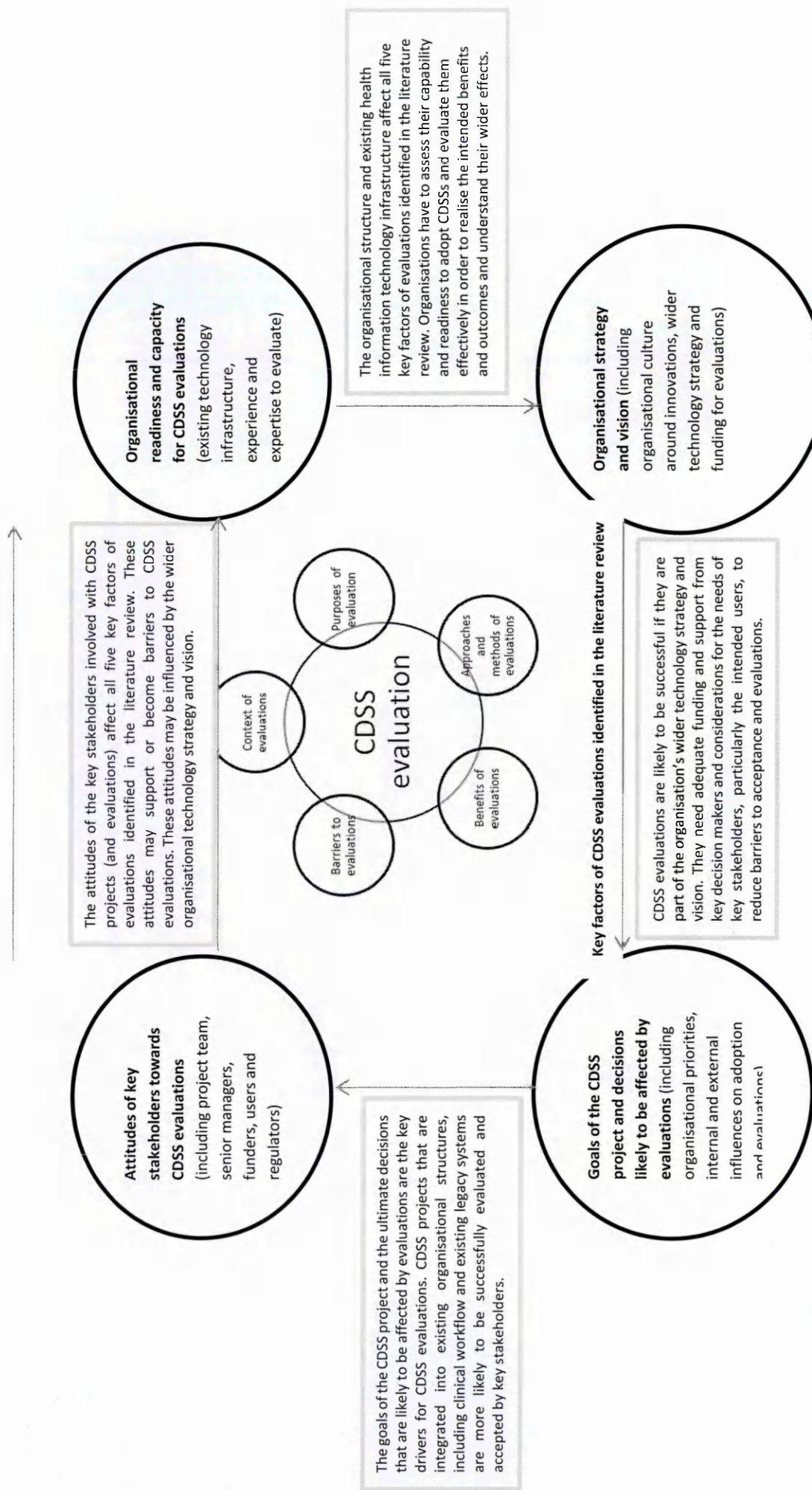


Figure 9.6 Synthesised CDSS evaluation framework

NHS Trusts are under pressure to implement computerised health information technologies (including CDSSs), to help them cope with increasing demand for health services and to support frontline healthcare workers with clinical decision making. It is imperative for NHS Trusts to incorporate these technologies into their wider organisational business strategies, technology strategies and organisational vision. This will help to create a culture within the organisation where CDSSs are implemented in a structured manner, rather than in response to external pressures and to solve immediate problems. Organisational culture has been shown to have significant effects on other organisational variables and employees (Lund, 2003). It has been argued that organisations should maintain a strong advantageous culture to help them to achieve their goals and stay competitive (Daraei, 2012). For NHS organisations, this can be achieved by providing funding for CDSS projects, including allocating adequate resources for their evaluation at every stage of the adoption cycle. Executive support will also be required, not only for the CDSS project teams, but for all the stakeholders who will be affected by the CDSS, including the users and where necessary, patients. Such an approach will help to build confidence in the CDSSs, as well as encouraging the necessary “buy in” from the key stakeholders. Additionally, this will also help to reduce barriers to CDSS evaluations. This research has shown that the CDSS’s adoption process is closely intertwined with its evaluations. It is therefore essential to clearly spell out the intended goals of the CDSS project, including the key decisions that will be affected by the CDSS and its related evaluations. These goals may be the key drivers for evaluations and lack of understanding is likely to affect the success of both the CDSS project and evaluations. Sharing the wider organisational vision and how the organisation expects to benefit from the CDSS will help to clarify the responsibilities and benefits to all key stakeholders. The literature review showed that CDSSs that are integrated with the existing clinical processes, workflow



and legacy systems are likely to be more accepted and used as intended. Interoperability with existing systems is also likely to encourage acceptance and ownership of both the CDSS and its related evaluations. Once a culture is developed where evaluations are valued, barriers to implementation, acceptance and usage of CDSSs may be minimised and more stakeholders may also engage in evaluative activity as part of their routine work processes. This will also help to widen the target audiences of CDSS evaluations and encourage the utilisation of evaluation results.

### **9.6.1 Summary of research contributions to CDSS evaluation literature**

This research made contributions to CDSS evaluation literature across all five factors of evaluation identified in the CDSS evaluation framework and allowed it to be extended as shown in Figure 9.6. The key research findings are as follows:

**The CDSS evaluation framework** - The CDSS evaluation framework developed in this study was based on an extensive and systematic literature review. This evaluation framework also informed the methodology that was used to undertake this research. Other evaluators and researchers could also utilise this framework for their evaluations and also to extend it further.

**Primary research in real clinical settings** – this study involved primary research in real clinical settings, looking at novel technologies that are likely to be adopted widely in the NHS. By carrying out primary research to understand how CDSSs were evaluated in real NHS settings, this research contributed to CDSS evaluation literature across all five factors identified in the CDSS evaluation framework (Figure 2.2). In addition, findings from the three case studies contributed to an extended CDSS evaluation framework (Figure 9.6). These findings and the updated CDSS evaluation framework provide an important addition to CDSS evaluation

literature. The three case studies could also be useful for other stakeholders, such as NHS Trusts and the research community in general. The researcher is expecting to publish research papers based on the CDSS evaluation framework and key findings from the three case studies.

**Context of CDSS evaluations** – this research has shown that the context of CDSS evaluations helps decision makers to understand the complex social systems in which the CDSSs are introduced, their effects on these systems and how the CDSSs also affect these systems. The evaluation context is heterogeneous, even where CDSSs are introduced in the same NHS hospital. As such, it is essential for evaluators to not only ask “what works”, but “what works for whom, in what circumstances, in what respect and how”?

**Purposes of CDSS evaluations** – purposes of CDSS evaluations should expand beyond clinical validation and technical efficacy. In fact, socio-technical aspects are as important as CDSS clinical and technical efficacy and could help to answer some of the questions raised by Kaplan and others regarding the poor adoption, acceptance and usage of CDSSs. Furthermore, the novelty of CDSSs in the NHS brings new risks and make them challenging to evaluate. To understand these risks and wider effects, evaluation purposes need to look beyond achievement of CDSS project objectives.

**Approaches and methods of CDSS evaluations** - The majority of CDSS systematic reviews carried out since the 1990s were inconclusive and sometimes drew cautious conclusions. This was mainly due to the reported lack of CDSS homogeneity and study heterogeneity, as well as methodological limitations that excluded non-experimental studies. Rather than carrying out the research using the same methods [positivist], and potentially getting the same results, this research sought to gain an understanding of CDSS evaluations within a real NHS clinical

settings. It sought to examine them as they are, rather than how they are assumed or ought to be.

This research has shown that CDSS evaluations take a predominantly narrow view and that there is need for a more systemic approach that looks at the wider CDSS effects beyond set objectives. CDSS evaluation methods should be determined by the purposes of evaluations and the corresponding decisions that need to be made, rather than default assumptions, personal and professional preferences. An “eclectic mix” of evaluation approaches and methods is required at different stages of CDSS lifecycle.

**Barriers to CDSS evaluations** – barriers that are well publicised in the literature include technical, methodological, organisational, social and human factors. This research had found that stakeholder attitudes significantly contribute to barriers to CDSS evaluations in NHS settings. Ethical barriers were also found to be an important factor. Evaluation barriers could be minimised by employing a wider range of evaluation approaches and methods to help negotiate the political and ethical dilemmas, gaps, imperfections.

**Benefits of CDSS evaluations** – the benefits of CDSS evaluations were primarily focused on regulatory requirements and financial incentives. However, these benefits were often based on anecdotal evidence and were not communicated to key stakeholders. Evaluation benefits were sometimes exaggerated to meet the organisations, reputational, regulatory and financial needs. There was also lack of alignment of evaluations to wider organisational goals, which resulted in low usage of evaluations for decision making.

**Organisational strategy and vision** – alignment of CDSS evaluations to wider organisational goals, strategy and vision may help to improve ‘buy in’ from intended users, attracting support

and funding from key stakeholders and improving overall patient and organisational outcomes.

**Attitudes of key stakeholders towards CDSS evaluations** – the attitudes of key stakeholders towards evaluations had an effect on when and how the evaluations were carried out and on the methods and approaches employed and the uses to which the evaluations were put.

**Organisational readiness and capacity for CDSS evaluations** – organisational readiness and capacity for CDSS evaluation needs to be assessed, taking into account existing technology infrastructure, experience and expertise of evaluators. Following the findings from the literature review, there was an expectation of an orderly approach to CDSS adoption and evaluations. However, this research found that evaluations were at best messy. In some cases, there were no evaluations at all, and little or no evaluation documentation. Although most of the systematic reviews suggested that RCTs and other experimental studies were the most commonly methods, this research found that CDSS evaluations were mainly informal, and based on anecdotal evidence and professional judgments of key individuals rather than actual evidence. This contrary to expectation given the fact that the majority of the evaluators and key stakeholders in the CDSS adoption process were physicians who are presumed to favour RCTs and other experimental methods.

**Goals of CDSS projects** – CDSS project goals should be aligned to wider organisational strategic decision making and communicated to all key stakeholders.

**Decisions that are likely to be affected by CDSS evaluations** – evaluations ultimately have an impact on professional, organisational reputation and resource allocation. As such, less favourable evaluation results are often ignored, while some purported benefits may not be the direct result of CDSS interventions.

## 9.7 Potential limitations of this research

The main potential limitation of this research study was that it was carried out in a single NHS Trust and thus the results may not be generalisable to other NHS hospitals. However, to mitigate this potential limitation, the study looked at three CDSSs that had different circumstances surrounding their adoption:

1. T1 was developed by a multi-disciplinary project team in the study Trust on the existing clinical results reporting system which was already being used by the targeted users
2. T2 was developed by third party organisations in partnership with another NHS Trust and other partners. It was not integrated with the study Trust's existing legacy systems following its adoption
3. T3 was developed by a third party organisation in partnership with another NHS Trust. It was initially piloted on a single ward at the study Trust and was heavily customised to the study Trust's clinical systems and integrated with the existing clinical results reporting system before Trust-wide implementation

It can be argued, however, that other NHS Trusts faced similar circumstances and that the findings from the study are transferrable and could provide invaluable learning to the wider healthcare sector. All three CDSSs were followed through the various contexts of evaluations throughout their respective adoption cycles. All the key stakeholders who were involved with the adoption decisions, development, implementation and evaluation of the CDSSs were interviewed. Additionally various documents that were related to the CDSSs were reviewed as part of the data collection and analysis processes. These included NICE guidelines and

related national policy documents; the study Trust's operational policies, procedures and clinical pathways; the study Trust's annual reports and quality accounts, and evaluations reports where available. There are many acute teaching NHS Trusts who share similarities with the study Trust and are faced with similar challenges regarding the adoption and evaluation of CDSSs. The in-depth nature of this study and the resultant framework of the evaluations provide opportunities for these NHS Trusts learn from the experiences of the study Trust to inform their own evaluations or adoption decisions.

## **9.8 Conclusions and recommendations for further research**

This research has assessed evaluations of three CDSSs that were adopted by one of the largest acute teaching NHS Trusts in the UK. It began with a systematic literature review of the wider evaluation literature, and in particular, CDSS evaluation literature. This involved published and unpublished academic literature, and grey literature, such as NICE guidelines and other government documents, Trust policies and clinical pathways, press releases, financial reports and quality accounts. The research adopted an interpretive case study approach, using Yin's (2009b) embedded case study approach. Biernarcki and Waldorf's (Biernacki and Waldorf, 1981) snowball sampling technique was used to identify the key stakeholders and interview participants, who in turn referred or recommended their peers and others who had been involved at the different adoption and evaluation phases of the CDSSs. The interviews were semi-structured and used open-ended questions to allow the researcher to explore new lines of enquiry while maintaining the research focus.

Eisenhardt's (1989) within case analysis was applied to gain an in-depth understanding and "intimate familiarity" of each CDSS as a separate entity to establish unique attributes and

patterns that were then followed up in the cross case analysis. The information gained from the interviews was corroborated with documentary evidence such as the study Trust's policies and clinical pathways and the NICE guidelines and other regulatory documents that supported them, evaluation reports and related artefacts where they were available, the study Trust's annual reports and quality accounts and other relevant internal and external documents. Site visits and observations were also carried out in the real clinical environments where the CDSSs were developed (in some cases) and where they were used to gain an understanding of the natural clinical environments. The site visits and observations further corroborated the interviews and documentary analysis. Cross case analysis involved presenting the data in different formats such as tables, lists, critical events and spreadsheets to identify similarities and differences across all three cases.

This research found that CDSS evaluations were mainly focused on showing the study Trust's adherence to NICE guidelines and providing information to key stakeholders such as the regulatory authorities and Commissioners. These evaluations were focused on assessing whether the respective CDSSs had achieved their stated objectives and did not look at their effects on the users, clinical workflow and patient outcomes. There were no causal links between some of the reported benefits and the CDSSs. The evaluations were mainly formative and informal. Most evaluations were based on the opinions and professional judgements of the respective CDSS project leaders as well as peer reviews, audits and before and after service evaluations. Evaluation documentation was minimal, and some key evaluations did not have any documentation at all. In contrast to the literature review, none of the evaluations used the RCT although the evaluators generally took a positivist stance. There were many missed opportunities to evaluate the wider effects of the CDSSs beyond

showing adherence to guidelines. Furthermore, this research has shown that CDSS evaluations take a predominantly narrow view and evidences the need for a more systemic approach. This research is important because CDSSs are likely to be widely adopted by NHS Trusts in future and it is imperative to improve our understanding of how we deploy and effectively evaluate them.

The following recommendations are made for further research:

- This research could be extended initially through a survey of NHS Trusts that have adopted similar CDSSs to test the generalisability of the findings.
- The CDSS evaluation framework that was developed in this research could be utilised to carry out exploratory studies in similar settings.
- More CDSS evaluations involving researchers and practitioners should be carried out using a mixed methods approaches.
- More research should be undertaken to define successes and failure of CDSS evaluations and also to establish the most effective evaluation methods that will help NHS organisations to achieve their intended benefits.





## 10. References

- Adlington K, Finn R, Ghafur S, et al. (2013) *Commissioning: what's the big deal*. Available at: <http://www.nhs.uk/NHSEngland/thenhs/about/Documents/Commissioning-FINAL-2015.pdf>.
- AHRQ. (2015) *What is comparative effectiveness?* Available at: <http://effectivehealthcare.ahrq.gov/index.cfm/what-is-comparative-effectiveness-research1/>.
- All-party parliamentary groups. (2015) *All-party parliamentary groups* Available at: [http://news.bbc.co.uk/democracylive/hi/guides/newsid\\_81000/81876.stm](http://news.bbc.co.uk/democracylive/hi/guides/newsid_81000/81876.stm)
- Ammenwerth E, Brender J, Nykanen P, et al. (2004) Visions and strategies to improved evaluation of health information systems. Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. *Journal of Medical Informatics* 73: 479-191.
- Ammenwerth E and De Keizer N. (2005) Trends in Evaluation Research 1982-2002. *Methods Inf Med* 44: 44-56.
- Anderson FA and Spencer FA. (2003) Risk factors for venous thromboembolism. *Circulation* 107: I-9-I-16.
- APPTG. (2013) *All-party parliamentary group Annual review* Available at: <http://apptg.org.uk/wp-content/uploads/2013/11/APPTG-Annual-Review-2013.pdf>
- Arkes HR, Shaffer VA and Medow MA. (2008) The influence of a physician's use of a diagnostic decision aid on the malpractice verdicts of mock jurors. *Medical Decision Making* 28: 201-208.

- Ashurst C and Doherty NF. (2003) Towards the formulation of a 'best practice' framework for benefits realisation in IT projects. *Electronic Journal of Information Systems Evaluation* 6: 1-10.
- Audit Commission. (1999) Critical to Success. The place of efficient and effective critical care services within the acute hospital. . London: Audit Commission Publications.
- Barnett GO, Cimino JJ, Hupp JA, et al. (1987) DXplain. An evolving diagnostic decision-support system. *JAMA* 258: 67-74.
- Bates DW, Cohen M, Leape LL, et al. (2001) Reducing the frequency of errors in medicine using information technology. *Journal of the American Medical Informatics Association* 8: 299-308.
- Bates DW and Gawande AA. (2003) Improving safety with information technology. *New England Journal of Medicine* 348: 2526-2534.
- Bath PA. (2008) Health informatics: current issues and challenges. *Journal of Information Science*.
- Beauchamp TL and Childress JF. (2001) *Principles of biomedical ethics*: Oxford university press.
- Berg M. (1999) Patient care information systems and healthcare work: a sociotechnical approach. *International Journal of Medical Informatics* 55: 87-101.
- Berlin A, Sorani M and Sim I. (2004) Characteristics of outpatient clinical decision support systems: a taxonomic description. *Studies in Health Technology and Informatics* 107: 578.
- Berlin A, Sorani M and Sim I. (2006) A taxonomic description of computer-based clinical decision support systems. *J Biomed Inform* 39: 656-667.

- Berner ES. (2007) *Clinical decision support systems: theory and practice*, New York: Springer.
- Biernacki P and Waldorf D. (1981) Snowball sampling: Problems and techniques of chain referral sampling. *Sociological methods & research* 10: 141-163.
- Black AD, Car J, Pagliari C, et al. (2011) The impact of eHealth on the quality and safety of healthcare: a systematic overview. *PLoS Medicine* 8: e1000387.
- Bleich HL. (1968) Computer evaluation of acid-base disorders. *Trans Assoc Am Physicians* 81: 184-189.
- Boland MR, Rusanov A, So Y, et al. (2014) From expert-derived user needs to user-perceived ease of use and usefulness: A two-phase mixed-methods evaluation framework. *Journal of biomedical informatics* 52: 141-150.
- Boonstra A and Broekhuis M. (2010) Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. *BMC Health Services Research* 10: 231.
- Bradley G. (2010) *Benefit Realisation Management: A practical guide to achieving benefits through change*: Gower Publishing, Ltd.
- Bray F, Lortet-Tieulent J, Ferlay J, et al. (2010) Prostate cancer incidence and mortality trends in 37 European countries: an overview. *European journal of cancer* 46: 3040-3052.
- Brewster D, Fraser L, Harris V, et al. (2000) Rising incidence of prostate cancer in Scotland: increased risk or increased detection? *BJU international* 85: 463-473.
- Bright TJ, Wong A, Dhurjati R, et al. (2012) Effect of Clinical Decision-Support Systems A Systematic Review. *Annals of Internal Medicine*.
- British Uro-oncology Group. (2009) MDT (Multi-disciplinary Team) Guidance for Managing Prostate Cancer

2nd Edition ed. London: British Uro-oncology Group and the BAUS Section of Oncology and the British Prostate Group.

British Uro-oncology Group. (2013) MDT (Multi-disciplinary Team) guidance for managing prostate cancer. London: British Uro-oncology Group and the BAUS Section of Oncology and the British Prostate Group.

Broadbent M, Weill P and Neo BS. (1999a) Strategic context and patterns of IT infrastructure capability. *The Journal of Strategic Information Systems* 8: 157-187.

Broadbent M, Weill P and St. Clair D. (1999b) The implications of information technology infrastructure for business process redesign. *MIS quarterly*: 159-182.

Brooks L, Atkinson C and Wainwright D. (2008) Adapting Structuration Theory to understand the role of reflexivity: Problematization, clinical audit and information systems. *International Journal of Information Management* 28: 453-460.

Buntin MB, Burke MF, Hoaglin MC, et al. (2011) The benefits of health information technology: a review of the recent literature shows predominantly positive results. *Health Affairs* 30: 464-471.

Burkle T, Ammenwerth E, Prokosch HU, et al. (2001) Evaluation of clinical information systems. What can be evaluated and what cannot? *J Eval Clin Pract* 7: 373-385.

Campbell DT. (1975) III. "Degrees of freedom" and the case Study. *Comparative political studies* 8: 178-193.

Campion Jr TR, Waitman LR, May AK, et al. (2010) Social, organizational, and contextual characteristics of clinical decision support systems for intensive insulin therapy: a literature review and case study. *International Journal of Medical Informatics* 79: 31-43.

- Cancer Research UK. (2014) *Prostate cancer incidence statistics*. Available at: <http://www.cancerresearchuk.org/cancer-info/cancerstats/types/prostate/incidence/uk-prostate-cancer-incidence-statistics>.
- Catwell L and Sheikh A. (2009) Evaluating eHealth interventions: the need for continuous systemic evaluation. *PLoS Medicine* 6: 830.
- Chaudhry B, Wang J, Wu S, et al. (2006) Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine* 144: 742-752.
- Chen HT. (1990) *Theory-driven evaluations*, Thousand Oaks, CA: Sage.
- Chen HT and Rossi PH. (1992) Using theory to improve program and policy evaluations. Santa Barbara, CA: Greenwood Press.
- Chung SH, Rainer Jr RK and Lewis BR. (2003) The impact of information technology infrastructure flexibility on strategic alignment and application implementations. *The Communications of the Association for Information Systems* 11: 44.
- Classen DC and Bates DW. (2011) Finding the Meaning in Meaningful Use. *The New England Journal of Medicine* 365: 855-858.
- Claxton K. (1999) The irrelevance of inference: a decision-making approach to the stochastic evaluation of healthcare technologies. *Journal of Health Economics* 18: 341-364.
- Cochrane AL. (1972) *Effectiveness and Efficiency: Random Reflections on Health Services*, London: The Royal Society of Medicine Press.
- Cohen AT, Tapson VF, Bergmann JF, et al. (2008) Venous thromboembolism risk and prophylaxis in the acute hospital care setting (ENDORSE study): a multinational cross-sectional study. *Lancet* 371: 387-394.

- Coiera E. (2003) Clinical decision support systems. *Guide to health informatics* 2.
- Coiera E, Westbrook J and Wyatt J. (2006) The safety and quality of decision support systems. *Methods Inf Med* 45: 20-25.
- Coldman AJ, Phillips N and Pickles TA. (2003) Trends in prostate cancer incidence and mortality: an analysis of mortality change by screening intensity. *Canadian Medical Association Journal* 168: 31-35.
- Collen MF, Rubin L, Neyman J, et al. (1964) Automated Multiphasic Screening and Diagnosis. *Am J Public Health Nations Health* 54: 741-750.
- Collin SM, Martin RM, Metcalfe C, et al. (2008) Prostate-cancer mortality in the USA and UK in 1975–2004: an ecological study. *The lancet oncology* 9: 445-452.
- Cooksey D. (2006) A Review of UK Health Research Funding. London: HMSO.
- Corbin JM and Strauss A. (1990) Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 13: 3-21.
- Cornford T and Klecun-Dabrowska E. (2001) Ethical perspectives in evaluation of telehealth. *Cambridge Quarterly of Healthcare Ethics* 10: 161-169.
- Coryn CLS, Noakes LA, Westine CD, et al. (2011) A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. *American Journal of Evaluation* 32: 199-226.
- Creswell JW. (2007) *Qualitative inquiry and research design: Choosing among five approaches: Choosing among five approaches*, Thousand Oaks, CA: Sage Publications, Inc.
- Cronbach LJ, Ambron SR, Dornbusch SM, et al. (1980) *Toward Reform of Program Evaluation*, San Francisco, CA: Jossey-Bass Publishers.
- Daraei JA. (2012) Relationship between organizational culture and strategy implementation: typologies and dimensions.

- Darling A. (1998) Modern Public Services for Britain: Investing in Reform. Comprehensive Spending Review: New Public Spending Plans 1999-2002. London: HM Treasury.
- Darzi A. (2007) NHS Next Stage Review Interim Report. London: Department of Health.
- de Dombal FT, Leaper DJ, Staniland JR, et al. (1972) Computer-aided diagnosis of acute abdominal pain. *British Medical Journal* 1: 9-13.
- Delpierre C, Cuzin L, Fillaux J, et al. (2004) A systematic review of computer-based patient record systems and quality of care: more randomized clinical trials or a broader approach? *International Journal for Quality in Healthcare* 16: 407-416.
- Denscombe M. (2003) *The good research guide for small-scale social research projects*, Maidenhead: Open University Press.
- Denvir MA and Leslie SJ. (2008) Best practice for chronic heart failure patients—writing guidelines is not enough. *European Heart Journal* 29: 1706-1708.
- Denzin NK and Lincoln YS. (2009) Qualitative research. *Yogyakarta: PustakaPelajar*.
- Department of Health. (1997) The new NHS: modern, dependable. London: Department of Health.
- Department of Health. (2000) Comprehensive critical care: a review of adult critical care services. London: The Department of Health.
- Department of Health. (2008) Treatment of Heart Attack National Guidance: Final Report of the National Infarct Angioplasty Project (NIAP). London: Department of Health.
- Department of Health. (2010) *Prevention of Venous Thromboembolism (VTE) in Hospitalised Patients*. Available at: <http://www.vteprevention-nhs.uk/images/ARCHIVE%20update%20on%20VTE%20prevention%20programme%202010.pdf>.
- Donabedian A. (2005) Evaluating the Quality of Medical Care. *Milbank Quarterly* 83: 691–729.



- Donaldson SI. (2007) *Program theory-driven evaluation science*, New York, NY: Lawrence Erlbaum.
- Douketis JD and Moinuddin I. (2008) Prophylaxis against venous thromboembolism in hospitalized medical patients: an evidence-based and practical approach. *Pol Arch Med Wewn* 118: 209-215.
- Dr Foster. (2011) *Guide rates best and worst hospitals in 2011*. Available at: <http://www.nhs.uk/news/2011/11November/Pages/2011-hospital-performance-data-analysed.aspx>.
- Dr Foster. (2015) *Dr Foster Intelligence*. Available at: <http://www.drfooster.com/> Accessed 25/08/2015.
- Drummond MF, Schwartz JS, Jönsson B, et al. (2008) Key principles for the improved conduct of health technology assessments for resource allocation decisions. *International journal of technology assessment in healthcare* 24: 244-258.
- Durieux P, Nizard R, Ravaud P, et al. (2000) A clinical decision support system for prevention of venous thromboembolism: effect on physician behavior. *JAMA* 283: 2816-2821.
- Eaton L, , . (2008) All myocardial infarction patients in England are to be offered angioplasty. *BMJ*: 2185.
- Eisenhardt KM. (1989) Building theories from case study research. *Academy of management review*: 532-550.
- Eisenstein EL, Kushniruk AW and Aarts J. (2011) What Works in User-Centered Health Informatics Evaluation? In: Moen A (ed) *23rd Conference of the European Federation for Medical Informatics*. MIE 2011.
- Eisner EW. (1979) *The Educational Imagination: On the Design and Evaluation of School Programs*, New York: Macmillan.

- Elliott D, Allen E, Perry L, et al. (2015) Clinical user experiences of observation and response charts: focus group findings of using a new format chart incorporating a track and trigger system. *BMJ quality & safety* 24: 65-75.
- Ellwood PM. (1988) Outcomes management. *New England Journal of Medicine* 318: 1549-1556.
- Eslami S, Abu-Hanna A, de Jonge E, et al. (2009) Tight glycemic control and computerized decision-support systems: a systematic review. *Intensive care medicine* 35: 1505-1517.
- Evans D. (2003) Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing* 12: 77-84.
- Farbey B, Land F and Targett D. (1993) How to evaluate your IT investment. Butterworth Heinemann, Oxford.
- Flyvbjerg B. (2011) Case Study. In: Denzin NK and Lincoln YS (eds) *The Sage Handbook of Qualitative Research*. Thousand Oaks, CA: Sage, 301-316
- Fournier DM. (1995) Establishing evaluative conclusions: A distinction between general and working logic. In: Fournier DM (ed) *Reasoning in evaluation: Inferential links and leaps*. San Francisco, CA: Jossey-Bass, 15-32.
- Fowler Jr FJ. (2013) *Survey research methods*: Sage publications.
- Franklin C and Mathew J. (1994) Developing strategies to prevent in-hospital cardiac arrest: analyzing responses of physicians and nurses in the hours before the event. *Critical care medicine* 22: 244-247.
- Friedman CP and Wyatt J. (2006) *Evaluation methods in biomedical informatics*: Springer Science & Business Media.

- Friedman CP and Wyatt JC. (1996) *Evaluation Methods in Medical Informatics*, New York: Springer.
- Gaba D and Howard SK. (1995) Situation awareness in anesthesiology. *Human Factors* 37: 20-31.
- Gao H, McDonnell A, Harrison DA, et al. (2007) Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intensive care medicine* 33: 667-679.
- Gardner-Thorpe J, Love N, Wrightson J, et al. (2006) The value of Modified Early Warning Score (MEWS) in surgical in-patients: a prospective observational study. *Annals of the Royal College of Surgeons of England* 88: 571.
- Garg AX, Adhikari NK, McDonald H, et al. (2005a) Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 293: 1223-1238.
- Garg AX, Adhikari NKJ, McDonald H, et al. (2005b) Effects of computerized clinical decision support systems on practitioner performance and patient outcomes. *JAMA: the journal of the American Medical Association* 293: 1223.
- Geerts WH, Bergqvist D, Pineo GF, et al. (2008) Prevention of venous thromboembolism: American College of Chest Physicians evidence-based clinical practice guidelines. *CHEST Journal* 133: 381S-453S.
- Geerts WH, Heit JA, Clagett GP, et al. (2001) Prevention of venous thromboembolism. *CHEST Journal* 119: 132S-175S.
- Geerts WH, Pineo GF, Heit JA, et al. (2004) Prevention of venous thromboembolism: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *CHEST Journal* 126: 338S-400S.

- Gersick CJ. (1989) Marking time: Predictable transitions in task groups. *Academy of Management journal* 32: 274-309.
- Gilliland F, Becker TM, Smith A, et al. (1994) Trends in prostate cancer incidence and mortality in New Mexico are consistent with an increase in effective screening. *Cancer Epidemiology Biomarkers & Prevention* 3: 105-111.
- Glynne P. (2007) Benefits management-changing the focus of delivery. *Association for Progress Management Yearbook 2006* 7: 45-49.
- GMC. (2015) *The General Medical Council*. Available at: <http://www.gmc-uk.org/> Accessed 25/08/2015.
- Goldhill D, Worthington L, Mulcahy A, et al. (1999) The patient-at-risk team: identifying and managing seriously ill ward patients. *ANAESTHESIA-LONDON*- 54: 853-860.
- Goldhill DR and Sumner A. (1998) Outcome of intensive care patients in a group of British intensive care units. *Critical care medicine* 26: 1337-1345.
- Goldstein MK, Coleman RW, Tu SW, et al. (2004) Translating research into practice: organizational issues in implementing automated decision support for hypertension in three medical centers. *Journal of the American Medical Informatics Association* 11: 368-376.
- Greenhalgh T. (2007) *Primary Healthcare: Theory and Practice*, Oxford: Blackwells.
- Greenhalgh T, Humphrey C, Hughes J, et al. (2009) How do you modernize a health service? A realist evaluation of whole-scale transformation in london. *Milbank Quarterly* 87: 391-416.
- Greenhalgh T and Russell J. (2010a) Why Do Evaluations of eHealth Programs Fail? An Alternative Set of Guiding Principles. *PLoS Med* 7: 1-5.

- Greenhalgh T and Russell J. (2010b) Why do evaluations of eHealth programs fail? An alternative set of guiding principles. *PLoS Medicine* 7: 1331.
- Greenhalgh T, Stramer K, Bratan T, et al. (2010) *The devil is in the detail: final report of the independent evaluation of the Summary Care Record and HealthSpace programmes*, London: University College London.
- Guba E and Lincoln Y. (1989) *Fourth Generation Evaluation*, Newbury Pak, CA: Sage.
- Guskey TR. (2000) *Evaluating professional development*, Corwin: Thousand Oaks, CA.
- Hallin B, Holm T, Seeman lodding H, et al. (2011) Clinical Decision Support Systems - CDSS effects on adherence to clinical guidelines by active 'on screen' computer-based clinical decision support systems.
- Hammond KR. (1998) Judgment and decision making in dynamic tasks. *Inform Decis Technol* 14: 3-14.
- Hammond NE, Spooner AJ, Barnett AG, et al. (2013) The effect of implementing a modified early warning scoring (MEWS) system on the adequacy of vital sign documentation. *Australian Critical Care* 26: 18-22.
- Harrison S and Wood B. (2000) Scientific-bureaucratic medicine and U.K. health policy. *Policy Stud Rev* 17: 25-42.
- Haux R, Ammenwerth E, Ter Burg W, et al. (2004) An international course on strategic information management for medical informatics students: aim, content, structure, and experiences. *International Journal of Medical Informatics* 73: 97-100.
- Health Select Committee. (2005) *The Prevention of Venous Thromboembolism in Hospitalised Patients*, London: HMSO.
- Heathfield H, Pitty D and Hanka R. (1998) Evaluating information technology in healthcare: barriers and challenges. *BMJ* 316: 1959.

- Hemens BJ, Holbrook A, Tonkin M, et al. (2011) Computerized clinical decision support systems for drug prescribing and management: a decision-maker-researcher partnership systematic review. *Implement Sci* 6: 89.
- Heselmans A, Van de Velde S, Donceel P, et al. (2009) Effectiveness of electronic guideline-based implementation systems in ambulatory care settings-a systematic review. *Implement Sci* 4: 82-82.
- Holle R and Zahlmann G. (1999) Evaluation of telemedical services. *Information Technology in Biomedicine, IEEE Transactions on* 3: 84-91.
- Holt DT, Helfrich CD, Hall CG, et al. (2010) Are you ready? How health professionals can comprehensively conceptualize readiness for change. *Journal of General Internal Medicine* 25: 50-55.
- Hsiao CJ, Hing E, Socey TC, et al. (2010) *Electronic medical record/electronic health record systems of office-based physicians: United States, 2009 and preliminary 2010 state estimates.*
- Hsing AW, Tsao L and Devesa SS. (2000) International trends and patterns of prostate cancer incidence and mortality. *International Journal of Cancer* 85: 60-67.
- Hunt DL, Haynes RB, Hanna SE, et al. (1998) Effects of computer-based clinical decision support systems on physician performance and patient outcomes. *JAMA: the journal of the American Medical Association* 280: 1339-1346.
- INAHTA. (2015) *Health Technology Assessment*. Available at: <http://www.inahta.org/hta-tools-resources/>.
- IQNARC. (2015) *Intensive Care National Audit and Research Centre* Available at: <https://www.icnarc.org/>

- Jaspers MW, Smeulers M, Vermeulen H, et al. (2011a) Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. *J Am Med Inform Assoc* 18: 327-334.
- Jaspers MWM, Smeulers M, Vermeulen H, et al. (2011b) Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. *Journal of the American Medical Informatics Association* 18: 327.
- Jefford, Stockler MR and Tattersall MHN. (2003) Outcomes research: what is it and why does it matter? *Internal Medicine Journal* 33: 110–118.
- Joint Committee on Standards for Educational Evaluation. (1994) *The program evaluation standards*, Thousand Oaks, CA: Sage.
- Kahn SR, Panju A, Geerts W, et al. (2007) Multicenter evaluation of the use of venous thromboembolism prophylaxis in acutely ill medical patients in Canada. *Thrombosis Research* 119: 145-155.
- Kaplan B. (2001a) Evaluating informatics applications--some alternative approaches: theory, social interactionism, and call for methodological pluralism. *International Journal of Medical Informatics* 64: 39-56.
- Kaplan B. (2001b) Evaluating informatics applications - clinical decision support systems literature review. *International Journal of Medical Informatics* 64: 15-37.
- Kaplan B and Harris-Salamone KD. (2009a) Health IT success and failure: recommendations from literature and an AMIA workshop. *Journal of the American Medical Informatics Association* 16: 291-299.
- Kaplan B and Harris-Salamone KD. (2009b) Health IT success and failure: recommendations from literature and an AMIA workshop. *J Am Med Inform Assoc* 16: 291-299.

- Kaplan B and Maxwell J. (2005) Qualitative research methods for evaluating computer information systems. *Evaluating the Organizational Impact of Healthcare Information Systems*: 30-55.
- Kaplan B and Shaw NT. (2004) Future directions in evaluation research: people, organizational, and social issues. *Methods Inf Med* 43: 215-231.
- Kaplan RS and Norton DP. (1995) Putting the balanced scorecard to work. *Performance measurement, management, and appraisal sourcebook* 66.
- Karsh B-T, Weinger MB, Abbott PA, et al. (2010) Health information technology: fallacies and sober realities. *Journal of the American Medical Informatics Association* 17: 617-623.
- Kastner M and Straus SE. (2008) Clinical decision support tools for osteoporosis disease management: a systematic review of randomized controlled trials. *Journal of General Internal Medicine* 23: 2095-2105.
- Kaushal R, CShojania KG and Bates DW. (2005) Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. *Archives of Internal medicine* 163: 1409-1416.
- Kawamoto K, Houlihan CA, Balas EA, et al. (2005) Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 330: 765.
- Kirkpatrick DL. (1994) *Evaluating Training Programs: The Four Levels*, San Francisco: Berrett-Koehler.
- Klecun E and Cornford T. (2005) A critical approach to evaluation. *European Journal of Information Systems* 14: 229-243.
- Klein G and Calderwood R. (1991) Decision models: some lessons from the field. *Trans Syst Man Cybern* 21: 1018-1026.



- Klein HK and Myers MD. (1999) A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS quarterly*: 67-93.
- Knaus WA, Wagner DP, Draper EA, et al. (1991) The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 100: 1619-1636.
- Kohn LT, Corrigan J and Donaldson MS. (2000a) *To err is human : building a safer health system*, Washington, D.C. ; [Great Britain]: National Academy Press.
- Kohn LT, Corrigan JM and Donaldson MS. (2000b) *To err is human: building a safer health system*. A report of the Committee on Quality of Healthcare in America, Institute of Medicine. Washington, DC: National Academy Press.
- Kolic I, Crane S, McCartney S, et al. (2015) Factors affecting response to National Early Warning Score (NEWS). *Resuscitation* 90: 85-90.
- Kong G, Xu D and Yang J. (2008) Clinical decision support systems: A review on knowledge representation and inference under uncertainties. *International Journal of Computational Intelligence Systems* 1: 159-167.
- Krakau I and Fabian C. (1999) Who needs all the information collected in computerized medical records? A computer crash shows that to ask the patient is often simpler and quicker. *Läkartidningen* 96: 4032.
- Kushner SI. (2002) The object of one's passion: engagement and community in democratic evaluation. *Evaluation Journal Australasia* 2: 16-22.
- Kushniruk AW. (2001) Analysis of complex decision-making processes in healthcare: Cognitive approaches to health informatics. *J Biomed Inform* 34: 365-376.
- Kvale S. (1996a) *Interviews : an introduction to qualitative research interviewing*, Thousand Oaks ; London: SAGE.

- Kvale S. (1996b) *Interviews: An introduction to qualitative research interviewing*, Thousand Oaks, California: SAGE Publications, Inc. .
- Land F. (2000) Evaluation in a socio-technical context. *Organizational and social perspectives on information technology*. Springer, 115-126.
- Lau F, Kuziemsky C, Price M, et al. (2010) A review on systematic reviews of health information system studies. *Journal of the American Medical Informatics Association* 17: 637.
- Leape LL and Berwick DM. (2005) Five years after To Err Is Human: what have we learned? *JAMA* 293: 2384-2390.
- Ledley RS and Lusted LB. (1959) Reasoning Foundations of Medical Diagnosis Symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science* 130: 9-21.
- Lehmannm H and Ohno-Machado L. (2011) Evaluation of informatics systems: beyond RCTs and beyond the hospital. *Journal of the American Medical Informatics Association* 18: 110-111.
- Leprohon J and Patel VL. (1995) Decision strategies in emergency telephone triage. *Medical Decision Making* 15: 240-253.
- Leys M. (2003) Healthcare policy: qualitative evidence and health technology assessment. *Health Policy* 65: 217-226.
- Liddell A, Ayling M and Reid G. (2011) *Innovation Health and Wealth, Accelerating Adoption and Diffusion in the NHS*. London: Department of Health.
- Liu JL and Wyatt JC. (2011a) The case for randomized controlled trials to assess the impact of clinical information systems. *Journal of the American Medical Informatics Association* 18: 173-180.

- Liu JL and Wyatt JC. (2011b) The case for randomized controlled trials to assess the impact of clinical information systems. *J Am Med Inform Assoc* 18: 173-180.
- Lluch M. (2011) Healthcare professionals' organisational barriers to health information technologies—A literature review. *International Journal of Medical Informatics* 80: 849-862.
- Lobach D, Sanders GD, Bright TJ, et al. (2012) Enabling healthcare decisionmaking through clinical decision support and knowledge management.
- Lomotan EA, Hoeksema LJ, Edmonds DE, et al. (2012) Evaluating the use of a computerized clinical decision support system for asthma by pediatric pulmonologists. *Int J Med Inform* 81.
- Lund DB. (2003) Organizational culture and job satisfaction. *Journal of business & industrial marketing* 18: 219-236.
- Maanen JV. (1988) Tales of the field: On writing ethnography. *Chicago Guides to Writing, Editing, and Publishing*.
- Main C, Moxham T, Wyatt JC, et al. (2010) Computerised decision support systems in order communication for diagnostic, screening or monitoring test ordering: systematic reviews of the effects and cost-effectiveness of systems.
- Marcilly R, Bernonville S, Riccioli C, et al. (2012) Patient safety-oriented usability testing: a pilot study. *Stud Health Technol Inform* 180: 368-372.
- Mark MM, Henry GT and Julnes G. (1998) A realist theory of evaluation practice. In: Henry GT, Julnes G and Mark MM (eds) *Realist evaluation: An emerging theory in support of practice*. San Francisco, CA: Jossey-Bass, 3–31.
- Martinsons M, Davison R and Tse D. (1999) The balanced scorecard: a foundation for the strategic management of information systems. *Decision support systems* 25: 71-88.

- Massey D, Aitken LM and Chaboyer W. (2009) What factors influence suboptimal ward care in the acutely ill ward patient? *Intensive and Critical Care Nursing* 25: 169-180.
- Mathieson SA. (2011) England's NHS teaching hospitals: a capital affair. *The Guardian*.
- May C, Harrison R, Finch T, et al. (2003) Understanding the normalization of telemedicine services through qualitative evaluation. *Journal of the American Medical Informatics Association* 10: 596-604.
- McArthur-Rouse F. (2001) Critical care outreach services and early warning scoring systems: a review of the literature. *Journal of advanced nursing* 36: 696-704.
- McNeill G and Bryden D. (2013) Do either early warning systems or emergency response teams improve hospital patient survival? A systematic review. *Resuscitation* 84: 1652-1667.
- McQuillan P, Pilkington S, Allan A, et al. (1998) Confidential inquiry into quality of care before admission to intensive care. *BMJ* 316: 1853-1858.
- Metzger J and MacDonald K. (2002) *Clinical decision support for the independent physician practice* Oakland, CA: California Healthcare Foundation.
- MHRA. (2014) *Medical device stand-alone software including apps*. Available at: <https://www.gov.uk/government/publications/medical-devices-software-applications-apps/medical-device-stand-alone-software-including-apps>.
- MHRA. (2015) *The Medicines and Healthcare products Regulatory Agency: What We Do*. Available at: <https://www.gov.uk/government/organisations/medicines-and-healthcare-products-regulatory-agency/about>.
- Miles MB and Huberman AM. (1984) Drawing valid meaning from qualitative data: Toward a shared craft. *Educational researcher*: 20-30.

- Miller PL. (1983) Critiquing anesthetic management: the "ATTENDING" computer system. *Anesthesiology* 58: 362-369.
- Miller PL. (1986) Extending computer-based critiquing to a new domain: ATTENDING, ESSENTIAL-ATTENDING, and VQ-ATTENDING. *Int J Clin Monit Comput* 2: 135-142.
- Miller RA. (1994) Medical diagnostic decision support systems—past, present, and future. *Journal of the American Medical Informatics Association* 1: 8.
- Miller RA and Masarie FE, Jr. (1990) The demise of the "Greek Oracle" model for medical diagnostic systems. *Methods Inf Med* 29: 1-2.
- Miller RA, Pople HE and Myers JD. (1982) Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine* 307: 468-476.
- Moehr JR, Schaafsma J, Anglin C, et al. (2006) Success factors for telehealth—a case study. *International Journal of Medical Informatics* 75: 755-763.
- Mohammed M, Hayton R, Clements G, et al. (2009) Improving accuracy and efficiency of early warning scores in acute care. *British Journal of Nursing* 18: 18-24.
- Montgomery AA and Fahey T. (1998) A systematic review of the use of computers in the management of hypertension. *Journal of epidemiology and community health* 52: 520-525.
- Moore AL, Dimitropoulou P, Lane A, et al. (2009) Population-based prostate-specific antigen testing in the UK leads to a stage migration of prostate cancer. *BJU international* 104: 1592-1598.
- Mumford E. (1983) *Designing Human Systems for New Technology: The ETHICS Method*, Manchester: Manchester Business School.

National Cancer Intelligence Network. (2009) Cancer Incidence and Survival by Major Ethnic Group. London: National Cancer Intelligence Network Coordinating Centre, 1-63.

National patient Safety Agency. (2007) *Recognising and responding appropriately to early signs of deterioration in hospitalised patients*. Available at: <http://www.nrls.npsa.nhs.uk/EasySiteWeb/getresource.axd?AssetID=60151>.

NCAA. (2015) *National cardiac Arrest Audit*. Available at: <https://www.resus.org.uk/research/national-cardiac-arrest-audit/> Accessed 25/08/2015.

Newman JV, E. . (2006) Discriminating customers, responsible patients, empowered users: Consumerism and the modernisation of healthcare. *Journal of Social Policy* 35: 193-209.

NHS Choices. (2015) *NHS Choices Website*. Available at: <http://www.nhs.uk/Pages/HomePage.aspx> Accessed 28/09/2015.

NHS Confederation. (2011) *The future's digital: Mental health and technology*. London: NHS Confederation.

NHS Confederation. (2015) *Key Statistics on the NHS*. Available at: <http://www.nhsconfed.org/resources/key-statistics-on-the-nhs>.

NHS SDO. (2006) An assessment of the clinical effectiveness, cost and viability of NHS General Practitioners with Special Interest (GPSI) services. In: Health Do (ed). London: National Co-ordinating Centre for NHS Service Delivery and Organisation Research and Development.

NICE. (2007) *Acutely ill patients in hospital: Recognition of and response to acute illness in adults in hospital*. In: Excellence NifHaC (ed). London: NOCE.

- NICE. (2008) Prostate cancer: diagnosis and treatment. *Prostate Cancer Guidance, CG58*.  
London: NICE.
- NICE. (2010a) *Venous thromboembolism in adults admitted to hospital: reducing the risk*.
- NICE. (2010b) *Venous thromboembolism prevention quality standard*. Available at:  
<https://www.nice.org.uk/guidance/qs3> Accessed 23/02/2014.
- NICE. (2014) *Prostate cancer: diagnosis and treatment*. Available at:  
<https://www.nice.org.uk/guidance/cg175> Accessed 25/03/2015.
- NICE. (2015) NICE Guidance: Evidenced-based information for health, public health and social care professionals.
- NICE CG175. (2014) Prostate cancer: diagnosis and treatment. In: Excellence NIfHaC (ed).  
London: NICE.
- Nieuwlaat R, Connolly SJ, Mackay JA, et al. (2011) Computerized clinical decision support systems for therapeutic drug monitoring and dosing: a decision-maker-researcher partnership systematic review. *Implement Sci* 6: 90-100.
- NPSA. (2007) Safer care for the acutely ill patient: learning from serious incident London: National patient Safety Agency.
- NPSA. (2014) *Root Cause Analysis (RCA) investigation*. Available at:  
<http://www.nrls.npsa.nhs.uk/resources/collections/root-cause-analysis/>.
- NPSA. (2015) *The National Patient Safety Agency*. Available at: <http://www.npsa.nhs.uk/>  
Accessed 25/08/2015.
- Nykanen P, Brender J, Ammenwerth E, et al. (2009) Introducing guidelines for good evaluation practice in health informatics. *Stud Health Technol Inform* 150: 958-962.

- Oakley A, Strange V, Bonell C, et al. (2006) Health services research: process evaluation in randomised controlled trials of complex interventions. *BMJ: British Medical Journal* 332: 413.
- Oliver SE, May MT and Gunnell D. (2001) International trends in prostate-cancer mortality in the "PSA era". *International Journal of Cancer* 92: 893-898.
- Orlikowski WJ and Baroudi JJ. (1991) Studying information technology in organizations: Research approaches and assumptions. *Information systems research* 2: 1-28.
- Osheroff JA, Teich JM, Middleton B, et al. (2007) A roadmap for national action on clinical decision support. *Journal of the American Medical Informatics Association* 14: 141-145.
- Ovretveit J. (1998) *Evaluating health interventions: an introduction to evaluation of health treatments, services, policies and organizational interventions*: McGraw-Hill International.
- Palvia SC, Sharma RS and Conrath DW. (2001) A socio-technical framework for quality assessment of computer information systems. *Industrial Management & Data Systems* 101: 237-251.
- Patel VL and Kaufman DR. (1998) Medical informatics and the science of cognition. *Journal of the American Medical Informatics Association* 5: 493-502.
- Patel VL, Kaufman DR and Arocha JF. (2002) Emerging paradigms of cognition in medical decision-making. *Journal of biomedical informatics* 35: 52-75.
- Patel VL, Kushniruk AW, Yang S, et al. (2000) Impact of a computer-based patient record system on data collection, knowledge organization, and reasoning. *J Am Med Inform Assoc* 7: 569-585.
- Patton QM. (1978) *Utilization-Focused Evaluation*, California: Sage Publications.



- Pawson R and Manzano-Santaella A. (2012) A realist evaluation workshop. *Evaluation* 18: 176-191.
- Pawson R and Tilley N. (1997) *Realistic evaluation*, London: Sage.
- Pearson SA, Moxey A, Robertson J, et al. (2009) Do computerised clinical decision support systems for prescribing change practice? A systematic review of the literature (1990-2007). *BMC Health Services Research* 9: 154.
- Prytherch DR, Smith GB, Schmidt PE, et al. (2010) ViEWS—towards a national early warning score for detecting adult inpatient deterioration. *Resuscitation* 81: 932-937.
- Quinn M and Babb P. (2002) Patterns and trends in prostate cancer incidence, survival, prevalence and mortality. Part I: international comparisons. *BJU international* 90: 162-173.
- Rahimi B and Vimarlund V. (2007) Methods to evaluate health information systems in healthcare settings: a literature review. *Journal of medical systems* 31: 397-432.
- Randell R, Mitchell N, Dowding D, et al. (2007) Effects of computerized decision support systems on nursing performance and patient outcomes: a systematic review. *Journal of Health Services and Policy* 12: 242-249.
- RCN. (2015) The Royal College of Nursing
- RCS. (2015) *The Royal College of Surgeons*. Available at: <https://www.rcseng.ac.uk/about>
- Reiss G. (2006) *Gower Handbook of programme management*: Gower Publishing, Ltd.
- Rigby M. (2001) Evaluation: 16 powerful reasons why not to do it-and 6 over-riding imperatives. *Studies in Health Technology and Informatics*: 1198-1202.
- Rigby M, Forsström J, Roberts R, et al. (2001) Verifying quality and safety in health informatics services. *BMJ* 323: 552-556.

- Robertson J, Walkom E, Pearson SA, et al. (2010) The impact of pharmacy computerised clinical decision support on prescribing, clinical and patient outcomes: a systematic review of the literature. *International Journal of Pharmacy Practice* 18: 69-87.
- Rogers EM. (1995) *Diffusion of innovations*, New York ; London: Free Press.
- Rogers P, Petrosino A, Hacsí T, et al. (2000a) Program Theory Evaluation: Practice, Promise and Problems. In: Rogers P, Petrosino A, Hacsí T, et al. (eds) *New Directions in Evaluation series*. San Francisco: CA: Jossey-Bass, 5–13.
- Rogers PJ, Petrosino A, Huebner TA, et al. (2000b) Program theory evaluation: Practice, promise, and problems. In: Rogers PJ, Hacsí TA, Petrosino A, et al. (eds) *Program theory in evaluation: Challenges and opportunities*. San Francisco, CA: Jossey-Bass, 5–14.
- Roper WL. (1988) The resource-based relative value scale: a methodological and policy evaluation. *JAMA* 260: 2444-2446.
- Roshanov PS, Misra S, Gerstein HC, et al. (2011) Computerized clinical decision support systems for chronic disease management: a decision-maker-researcher partnership systematic review. *Implement Sci* 6: 92.
- Rouse D. (2011) Employing Kirkpatrick's Evaluation Framework to Determine the Effectiveness of Health Information Management Courses and Programs. *Perspectives in Health Information Management* 8: 1-5.
- Rowan K, Kerr J, Major E, et al. (1993) Intensive Care Society's APACHE II study in Britain and Ireland--II: Outcome comparisons of intensive care units after adjustment for case mix by the American APACHE II method. *BMJ* 307: 977-981.

- Royal College of Physicians. (2012) National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. London: The Royal College of Physicians.
- Sahota N, Lloyd R, Ramakrishna A, et al. (2011) Computerized clinical decision support systems for acute care management: a decision-maker-researcher partnership systematic review of effects on process of care and patient outcomes. *Implement Sci* 6: 91-98.
- Sapountzis S, Harris K and Kagioglou M. (2008) Benefits Management and Benefits Realisation, a Literature Report. *HaCIRIC, april*.
- Sax FL and Charlson ME. (1987) Medical patients at high risk for catastrophic deterioration. *Critical care medicine* 15: 510-515.
- Savory, Clive and Fortune, Joyce (2013). *NHS adoption of NHS-developed technologies*. NIHR Service Delivery and Organisation Programme, Southampton, UK.
- Schedlbauer A, Prasad V, Mulvaney C, et al. (2009) What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior? *Journal of the American Medical Informatics Association* 16: 531-538.
- Schein R, Hazday N, Pena M, et al. (1990) Clinical antecedents to in-hospital cardiopulmonary arrest. *CHEST Journal* 98: 1388-1392.
- Schwartz WB and Mendelson DN. (1991) Hospital cost containment in the 1980s: hard lessons learned and prospects for the 1990s. *New England Journal of Medicine* 324: 1037-1042.
- Scriven M. (1972) *The methodology of evaluation. Perspectives of curriculum evaluation*, Chicago: Rand McNally.

- Shiffman RN, Freudigman KA, Brandt CA, et al. (2000) A guideline implementation system using handheld computers for office management of asthma: effects on adherence and patient outcomes. *Pediatrics* 105: 767-773.
- Shiffman RN, Liaw Y, Brandt CA, et al. (1999) Computer-based guideline implementation systems. *Journal of the American Medical Informatics Association* 6: 104-114.
- Shojania KG, Jennings A, Mayhew A, et al. (2009) The effects of on-screen, point of care computer reminders on processes and outcomes of care. *Cochrane Database Syst Rev* 3.
- Shortliffe EH, Davis R, Axline SG, et al. (1975) Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system. *Computers and biomedical research* 8: 303-320.
- Sim I, Gorman P, Greenes RA, et al. (2001) Clinical decision support systems for the practice of evidence-based medicine. *J Am Med Inform Assoc* 8: 527-534.
- Sintchenko V, Magrabi F and Tipper S. (2007) Are we measuring the right end-points? Variables that affect the impact of computerised decision support on patient outcomes: a systematic review. *Informatics for Health and Social Care* 32: 225-240.
- Sittig DF. (1994) Grand challenges in medical informatics? *Journal of the American Medical Informatics Association* 1: 412.
- Smith AF and Wood J. (1998) Can some in-hospital cardio-respiratory arrests be prevented? A prospective survey. *Resuscitation* 37: 133-137.
- Smith GB, Prytherch DR, Schmidt P, et al. (2006) Hospital-wide physiological surveillance—a new approach to the early identification and management of the sick patient. *Resuscitation* 71: 19-28.

- Smith GB, Prytherch DR, Schmidt PE, et al. (2008a) Review and performance evaluation of aggregate weighted 'track and trigger' systems. *Resuscitation* 77: 170-179.
- Smith GB, Prytherch DR, Schmidt PE, et al. (2008b) A review, and performance evaluation, of single-parameter "track and trigger" systems. *Resuscitation* 79: 11-21.
- Smith NL. (1993) Improving evaluation theory through the empirical study of evaluation practice. *Evaluation Practice* 14: 237-242.
- Smith R. (1992) Using a mock trial to make a difficult clinical decision. *BMJ* 305: 1284-1287.
- Smith R. (1998) All changed, changed utterly. *BMJ* 316: 1917-1918.
- Smithson S and Hirschheim R. (1998) Analysing information systems evaluation: another look at an old problem. *European Journal of Information Systems* 7: 158-174.
- Snyder RA and Fields WL. (2006) Measuring hospital readiness for information technology (IT) innovation: A multisite study of the Organizational Information Technology Innovation Readiness Scale. *Journal of Nursing Measurement* 14: 45-55.
- Stake RE. (1975) *Evaluating the Arts in Education: A Responsive Approach*, Columbus, OH: Merrill.
- Stelfox HT, Palmisani S, Scurlock C, et al. (2006) The "To Err is Human" report and the patient safety literature. *Quality and Safety in Healthcare* 15: 174-178.
- Stufflebeam DL. (1973a) Evaluation as enlightenment for decision-making. In: Worthen BR and Sanders JR (eds) *Educational evaluation: Theory and practice*. Worthington, OH: Charles A. Jones Publishing Company.
- Stufflebeam DL. (1973b) *Toward a science of educational evaluation*, Englewood Cliffs, NJ: Educational Technology Publications.

- Stufflebeam DL and Shinkfield AJ. (2007) *Evaluation theory, models, and applications*, San Francisco: Jossey-Bass.
- Subbe C, Davies R, Williams E, et al. (2003) Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions\*. *Anaesthesia* 58: 797-802.
- Subbe CP, Kruger M, Rutherford P, et al. (2001) Validation of a modified Early Warning Score in medical admissions. *QJM* 94: 521-526.
- Suchman E. (1967) *Evaluative research: Principles and practice in public service and social action programs*, New York: Russell Sage.
- Symons VJ. (1991) A review of information systems evaluation: content, context and process. *European Journal of Information Systems* 1: 205-212.
- Talmon J, Enning J, Castañeda G, et al. (1999) The VATAM guidelines. *International Journal of Medical Informatics* 56: 107-115.
- Tan K, Dear PR and Newell SJ. (2005) Clinical decision support systems for neonatal care. *The Cochrane Library*.
- Teasdale G. (2002) Learning from Bristol: report of the public inquiry into children's heart surgery at Bristol Royal Infirmary 1984-1995. *British journal of neurosurgery* 16: 211-216.
- Teich JM. (1999) Inpatient order management. *Journal of Healthcare Information Management* 13: 97-110.
- The Academy of Royal Colleges. (2015) *The Academy of Royal Colleges* Available at: <http://www.aomrc.org.uk/>.

The BBC. (2013) *The changing NHS*. Available at: <http://www.bbc.co.uk/news/health-19674838>.

The British Society for Haematology. (2015) *The British Society for Haematology* Available at: <http://www.b-s-h.org.uk/>

The British Society of Urological Surgeons. (2015) *The British Society of Urological Surgeons*. Available at: <http://www.baus.org.uk/>

The Department of Health. (2010) *Equity and excellence: Liberating the NHS*. In: Health TDo (ed). London: The Stationary Office Ltd.

The Department of Health. (2011) *Health and Social Bill*. In: Health TDo (ed). London.

The Department of Health. (2014) *The health and care system explained*. Available at: <https://www.gov.uk/government/publications/the-health-and-care-system-explained/the-health-and-care-system-explained>.

The Health and Social Care Information Centre. (2015) *The Health and Social Care Information Centre* Available at: <http://webarchive.nationalarchives.gov.uk/20130502102046/http://www.hscic.gov.uk/home>

The Law Commission. (2014) *Regulation of Health and Social Care Professionals*. Available at: <https://www.gov.uk/government/publications/regulation-of-health-and-social-care-professionals>

The Nursing and Midwifery Council. (2015) *The Nursing and Midwifery Council* Available at: <http://www.nmc-uk.org/>.

The Royal College of Midwives. (2015) *The Royal College of Midwives* Available at: <https://www.rcm.org.uk/>

The Royal College of Obstetricians and Gynaecologists. (2015) The Royal College of Obstetricians and Gynaecologists.

The Royal College of Physicians. (2015) *The Royal College of Physicians*. Available at: <https://www.rcplondon.ac.uk/>.

Tierney WM, Miller ME and McDonald CJ. (1990) The effect on test ordering of informing physicians of the charges for outpatient diagnostic tests. *New England Journal of Medicine* 322: 1499-1504.

Triggle N. (2014) NHS death rates 'should be ignored'.

Trivedi M, Kern J, Marcee A, et al. (2002) Development and implementation of computerized clinical guidelines: barriers and solutions. *Methods of Information in Medicine* 41: 435-442.

Tyler RW. (1942) General statement on evaluation. *Journal of Educational Research* 35: 492–501.

van Gemert-Pijnen J, Nijland N, van Limburg M, et al. (2011) A Holistic Framework to Improve the Uptake and Impact of eHealth Technologies. *Journal of Medical Internet Research* 13.

van Teijlingen E and Hundley V. (2001) The importance of pilot studies. *Social research update*: 1-4.

Vardi M and Haran M. (2012) Venous thromboembolism prophylaxis of acutely ill hospitalized medical patients. Are we over-treating our patients? *European journal of internal medicine* 23: 231-235.

VTE England. (2012) *VTE England Programme Timeline*. Available at: <http://www.vteprevention-nhsengland.org.uk/programme-timeline>.



- Walsham G. (1995) Interpretive case studies in IS research: nature and method. *European Journal of Information Systems* 4: 74-81.
- Walsham G. (2006) Doing interpretive research. *European Journal of Information Systems* 15: 320-330.
- Wanless D. (2002) *Securing our Future Health: Taking a Long-Term View*. London: HM Treasury.
- Ward J and Daniel E. (2006) *Benefits management: Delivering value from IS & IT investments*: John Wiley & Sons Chichester.
- Ward J and Elvin R. (1999) A new framework for managing IT-enabled business change. *Information systems journal* 9: 197-221.
- Warner HR, Toronto AF, Veasey LG, et al. (1961) A mathematical approach to medical diagnosis. Application to congenital heart disease. *JAMA* 177: 177-183.
- Weiner BJ. (2009) A theory of organizational readiness for change. *Implement Sci* 4: 67.
- Weiss CH. (1972) *Evaluation Research: Methods of Assessing Program Effectiveness*, Englewood Cliffs, N.J: Prentice-Hall.
- Weiss CH. (1995) Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In: Connell JA, Kubisch A, Schorr LB, et al. (eds) *New approaches to evaluating community initiatives: Volume 1, concepts, methods, and contexts*. New York, NY: Aspen Institute, 65-92.
- Weiss CH. (1997) Theory-based evaluation: Past, present, and future. *New Directions for Evaluation* 1997: 41-55.

- Weiss CH. (1998) *Evaluation: Methods for studying programs and policies (2nd ed.)*, Upper Saddle River, NJ: Prentice Hall.
- Weiss CH. (2004) On theory-based evaluation: Winning friends and influencing people. *The Evaluation Exchange* IX: 1-5.
- Wheeler-Carmichael G. (2000) GOVERNMENT IT PROJECTS: THE MCCARTNEY REPORT "SUCCESSFUL IT: MODERNIZING GOVERNMENT IN ACTION" AND THE CSSA REPORT "GETTING IT RIGHT FOR GOVERNMENT". *Computer Law & Security Review* 16: 325-328.
- Winter A, Ammenwerth E, Bott OJ, et al. (2001) Strategic information management plans: the basis for systematic information management in hospitals. *International Journal of Medical Informatics* 64: 99-109.
- Wolfstadt JI, Gurwitz JH, Field TS, et al. (2008) The effect of computerized physician order entry with clinical decision support on the rates of adverse drug events: a systematic review. *Journal of General Internal Medicine* 23: 451-458.
- Wyatt J. (1994a) Clinical data systems, part 1: data and medical records. *The Lancet* 344: 1543-1547.
- Wyatt J. (1994b) Clinical data systems, part 2: components and techniques. *The Lancet* 344: 1609-1614.
- Wyatt J and Spiegelhalter D. (1990) Evaluating medical expert systems: what to test and how? *Informatics for Health and Social Care* 15: 205-217.
- Wyatt J and Spiegelhalter D. (1991) Field trials of medical decision-aids: potential problems and solutions. *Proceedings of the annual symposium on computer application in medical care*. American Medical Informatics Association, 3.

- Wyatt JC. (1994c) Clinical data systems, part 3: development and evaluation. *The Lancet* 344: 1682-1688.
- Wyatt JC and Keen J. (2001) The new NHS information technology strategy: Technology will change practice. *BMJ: British Medical Journal* 322: 1378.
- Wyatt JC and Wyatt SM. (2003) When and how to evaluate health information systems? *International Journal of Medical Informatics* 69: 251-259.
- Yin RK. (2009a) *Case study research: design and methods*, London: Sage.
- Yin RK. (2009b) *Case study research: Design and methods*: Sage publications, INC.
- Yourman L, Concato J and Agostini JV. (2008) Use of computer decision support interventions to improve medication prescribing in older adults: a systematic review. *American Journal of Geriatric Pharmacotherapy* 6: 119-129.
- Zoritch B, Roberts I and Oakley A. (1998) The health and welfare effects of day-care: a systematic review of randomised controlled trials. *Social Science & Medicine* 47: 317-327.

# Appendices

## Appendix 1 - List of key Study Trust documents and secondary materials

### Appendix 1.1 - General study Trust documents

1. Study Trust Quality Accounts (2010/11)
2. Study Trust Quality Accounts (2011/12)
3. Study Trust Quality Accounts CQUIN Supplement Quality Account (2011/12)
4. Study Trust/Strategic Health Authority CQUIN Scheme Agreement (2011/12)
5. Strategic Commissioning Group - Specialised Commissioning Team meeting minutes (2009)
6. Award winning Trust challenges Dr Foster rating - press release (2009)
7. Study Trust clinical results reporting systems
8. Study Trust ethics approval

### Appendix 1.2 - T1 study Trust documents & secondary materials

1. Responsibility of Emergency Department and Ward staff during VTE RA BCP Activation
2. Responsibility of Emergency Department and Ward staff during VTE RA BCP De-activation
3. Clexane – summary of product characteristics
4. Copy of DH VTE Reporting graphs static 2010 – 2011
5. Clinical results reporting system VTE risk assessment function for doctors
6. Clinical results reporting system VTE risk assessment function for nursing staff (non-midwives)
7. Clinical results reporting system VTE risk assessment function for midwives
8. Department of Health VTE Reporting graphs static 2011 – 2012
9. DH VTE Reporting graphs static 2012 – 2013
10. Displaying the VTE RA Statuses and Viewing a VTE RA Form
11. Study Trust VTE RA Process – ED Admission
12. Study Trust VTE RA Process – Elective Admission
13. Study Trust VTE RA Process – Emergency Admission (straight to theatre)
14. Study Trust VTE RA Process – Theatres (Elective admission with pre-assessment)
15. Study Trust VTE RA Process – Theatres (Elective admission without pre-assessment)
16. Study Trust Venous Thromboembolism (VTE) Operational Policy
17. Strategic Health Authority letter authorising the exemption of day surgery patient cohort
18. Study Trust care pathway for adults with suspected DVT
19. Study Trust associated records
20. Study Trust Prescribing, Handling, Custody and Administration of Drugs Policy
21. Study Trust Patient Information Procedure UHCW NHS Trust ICT Security Policy
22. Study Trust Warfarin Guideline
23. Study Trust Heparin Treatment Guideline
24. An electronic tool to help prevent potentially fatal blood clots associated with hospitalisation (Submission to NICE shared learning database, 2011)
25. T1 portal

26. T1 PDA devices
27. T1 intranet updates

#### **Appendix 1.3 - T2 study Trust documents & secondary materials**

1. T2 portal
2. T2 clinical pathways
3. T2 GP outcome letter template
4. T2 developer's website
5. T2 promotional video 1
6. T2 promotional video 2
7. T2 patient survey

#### **Appendix 1.4 - T3 study Trust documents & secondary materials**

1. Frequently Asked Questions and Troubleshooting Guide for T3 (May 2008) - T3 FAQs
2. T3 New Ward Profile Checks (checklist)
3. T3 Training of Staff (checklist)
4. T3 Clinical Documents and process checks (checklist)
5. T3 Environmental analysis Checks (checklist)
6. Critical Care Outreach Activity Report 1<sup>st</sup> April 2012 – 31<sup>st</sup> March 2013
7. T3 Ward: Going Paperless Pilot Checklist
8. T3 Print from PDA Profile Test for Ward
9. Printer Profile for Ward
10. T3 Operational Policy August 2012
11. T3 presentation – “a means of hospital-wide physiological surveillance” (2009)
12. T3 Freedom of Information Request (March 2013)
13. T3 press release

## Appendix 2 - General indicative interview topics and questions

<p><b>Opening remarks</b> - Introduction, general overview of interview format, audio recording, and ethical issues and obtain brief background information about the participant</p> <p><b>Background of the CDSS</b></p> <ol style="list-style-type: none"> <li>When did you first hear about the CDSS? <ul style="list-style-type: none"> <li>Within or outside the Trust?</li> </ul> </li> <li>Were you already aware of the CDSS or did you have to obtain more information?</li> <li>What were the key considerations for choosing this CDSS? <ul style="list-style-type: none"> <li>Previously undertaken evaluations, potential benefits, Trust, collegial or national guidelines, peer networks, commercial organisations etc.</li> <li>What were the intended purposes of the CDSS? – Clinical, economic outcomes etc. and note anticipated benefits</li> <li>How was the CDSS brought to the Trust? – Prompt re: by whom and why?</li> <li>Was a business case prepared? Is there documentary evidence?</li> </ul> </li> </ol>
<p><b>Participant's involvement in the project</b></p> <ol style="list-style-type: none"> <li>How was the CDSS implemented? <ul style="list-style-type: none"> <li>Pilot, phased or trust-wide implementation</li> <li>Who did what, when and how? Level of training provided</li> </ul> </li> <li>What were the benefits and challenges during and after implementation? <ul style="list-style-type: none"> <li>Is there documented evidence of benefits e.g., quality improvement, effectiveness, cost savings, patient outcomes etc.</li> <li>Is there evidence of drawbacks or challenges? How were they reported? To whom? Was there a feedback system?</li> <li>How were these addressed and by whom? – is there evidence?</li> <li>Was there adequate support during and post implementation?</li> </ul> </li> <li>What was the effect of the CDSS on clinical workflow? <ul style="list-style-type: none"> <li>Were there any issues with usability/interface of the system?</li> <li>Degree of integration with existing systems and work processes</li> <li>Effect on clinical decision-making and patient outcomes</li> </ul> </li> </ol>
<p><b>Participant's involvement in the evaluation</b></p> <ol style="list-style-type: none"> <li>Were anticipated benefits realised? <ul style="list-style-type: none"> <li>At personal, team, departmental or hospital level</li> <li>Was success failure of the system clearly (pre)-defined? By whom and what were the parameters?</li> <li>Was there a communication system between the implementation team, trainers and frontline staff? E.g., feedback system, regular updates</li> <li>Were there any departmental/staff level meetings to discuss the implementation or respond to experiences in clinical settings?</li> <li>Who attended these meetings?</li> <li>What was the structure for reporting the project and evaluation? – is there documentary evidence? Internal/organisational use? External use?</li> </ul> </li> </ol>
<p><b>Stakeholder perceptions about the project and evaluation</b></p> <ol style="list-style-type: none"> <li>What were the key lessons from the implementation and evaluation? <ul style="list-style-type: none"> <li>Would any lessons learnt affect CDSS or other healthcare technology evaluation?</li> <li>Establish whether they consider their involvement/activities as evaluation?</li> </ul> </li> </ol>
<p><b>Closing remarks</b></p>

## Appendix 3 - Guide to evaluation inputs, processes and outcomes

### Evaluation inputs

1. Resources allocated e.g., budget of the evaluation and other material resources available to evaluators
2. Duration (s) of evaluation (s)
3. Timing/intervals of evaluation in relation to the entire project
4. Were evaluations undertaken by project personnel or dedicated team/individual (s)?
  - Who conducted the evaluation (s)? How were they selected, by whom?
  - Characteristics of evaluation personnel e.g., training, experience, commitment , 'world view' and attitudes towards evaluation
  - Characteristics of project personnel e.g., training, experience and 'world view' e.g., training, experience, commitment , 'world view' and attitudes towards evaluation
5. Design of methodologies used in evaluation
6. Characteristics of the project being evaluated, e.g., goals, substantive area of concern, client group, targeted users of the systems
7. Audience or client group or decision makers for the evaluation and purposes of evaluation - targeted decision makers
  - Uses of information generated by evaluation (s), e.g., decision-making about (re) allocation of resources
8. The evaluation setting
  - Was there a pre-determined criteria for success or failure?
  - Did the tool achieve the intended goals?
9. What were the key milestones, successes, drawbacks, feedback system (from evaluation to the project), modifications to evaluation or the project itself?

### Evaluation process

#### Transforming inputs into outcomes – actual decisions made

- 1 Types, intensity and frequency of interactions between evaluators and program staff members - may depend on whether CDSS evaluations are considered to be self-evaluations, or if independent department or third party are given the evaluation responsibility
- 2 Response of program staff and client groups to the presence of evaluators – Hawthorne effect (inadvertent response to evaluators by programme staff and users) – may depend on how evaluation was undertaken – in this case may not be an issue as above, especially within the same organisation – Was the Hawthorne effect considered and what efforts were made to minimise or address it?
- 3 Extent to which information acquired during the evaluation is fed back to program staff, perhaps modifying the program procedures – maybe deliberately to influence operation, or whether evaluators are actively involved or outsiders.
- 4 Extent to which information acquired during the evaluation is used to modify the allocation of evaluation resources – and project resources – how much is emerging info used to modify the evaluation of the project itself or alternative hypothesis or other uses of the CDSS (adaptability)?

- 5 Adaptiveness of the evaluation design (i.e., capacity to respond to changes in the program) and history of adaptations – deliberate or forced adaptations due to internal or external pressures and unintended effects of such activities – to what extent is the project jeopardised or able to adapt to such changes?
- 6 Turnovers in key personnel (e.g., evaluators, program staff, client groups of both program and evaluation) – effect of such changes on the project and evaluation, and the effect on the “decision impact” of the outcomes – note the “vanishing advocate”
- 7 Testing of hypothesis regarding the program – statistical correctness, while important, may not guarantee “decision impact” and statistically imperfect information may affect decision-making, which may be costly to the organisation. This may depend on the dominant school of thought.
- 8 Documentation of findings (in the case of SQUIN and NICE – what are the national averages and how do local results compare?). Evaluation report – is it part of the evaluation process or an outcome of evaluation? The report could also be viewed as communication of final evaluation findings, together with oral presentations, teaching and teaching aids to convey essential results, or any other activities to communicate or disseminate results.

#### Evaluation outcomes

- 1 Who is the primary group of evaluation – who is evaluating?
  - How close are they to project staff?
- 2 What are the ultimate decision consequences? - Possible decisions influenced by the evaluation:
  - Decision by the project commissioners or organisation to fund, continue to fund, modify or cancel a project or effect on such future decisions (or maintain the status quo)
  - Decision by the project personnel to modify any of the project procedures
  - Decision by users to alter system use
  - Decisions by internal or external staff to study further questions or issues raised in the evaluation – may be difficult to assess
  - Decision by other organisations, commissioners or similar project personnel to introduce, continue to fund, modify or terminate similar projects based on information from the evaluation - may be difficult to assess
  - There may be challenges in separating process and outcome (measures)



**Appendix 4 - Participant Information Sheet**



**Faculty of Mathematics, Computing and Technology**  
Department of Communication and Systems  
The Open University  
Walton Hall  
Milton Keynes  
United Kingdom  
MK7 6AA  
Tel +44 (0) 1908 653964

Participant Information Sheet Version 3 - 22/11/2012

**Study title – Evaluation of Clinical Decision-Support Technology in the NHS**

I am a PhD student at The Open University and I would like to invite you to take part in my research study looking at the evaluation of clinical decision-support technology in the NHS.

**Invitation to take part in a research study**

You are being invited to take part in this research study because you have been involved in the implementation, use or evaluation of clinical decision-support technology within the hospital. Before you decide whether or not to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and ask any questions you may have. Thank you for taking time to consider this study.

**What is the purpose of the research study?**

The purpose of this study is to investigate how innovative clinical decision-support technology is evaluated in NHS clinical settings. The study will also assess the effect of evaluation on subsequent adoption decisions and diffusion of healthcare technology within your hospital and wider healthcare settings.

**What is involved in participating in this study?**

It is up to you to decide whether or not to take participate in this study and you should not feel pressurised into participating by others. If you wish to participate please sign the informed consent form in duplicate and wet ink. You will then be given a copy for your records and I will retain the second copy. After consenting, you will be asked to participate in a semi-structured interview lasting no more than one hour at a time convenient for you. Where possible interviews will be conducted face to face, but if this is not possible, the interview can be conducted by telephone. During the interview you will be asked about your experience of clinical decision-support technologies that have been adopted by the hospital. It is intended

that the interviews will be recorded to ensure concise data collection and you will be given an opportunity to verify, delete or add to and sign interview summaries (or confirm by email) before they are used to develop published outputs and the PhD thesis. Please state on the consent form if you do not wish your interview to be audio recorded and notes will be taken instead. It may be necessary to contact you again to obtain clarification, ask further questions or request for relevant documentary data. You are free to withdraw from the study at any time without giving a reason (before the results are published) and it is intended that your data collected up to that point would be used for the study. Please state on the consent form if you do not wish your data to be used in the event that you withdraw.

#### **What are the possible benefits and risks of taking part?**

Taking part in this study gives you the opportunity to reflect on your practice and contribute your views and experiences on how healthcare technologies are evaluated. Your contributions will be used to develop a PhD thesis and associated published outputs, which will add to knowledge in this important area and inform future adoption and evaluation of healthcare technology, as well as decision-making in the NHS and the wider healthcare context. Published outputs include journal articles, peer discussions and conference presentations. There are no significant risks envisaged but I will undertake to resolve any concerns or potential risks arising in the course of the research. The Open University Human Research Ethics Committee (HREC) (Ref: HREC/2012/#1213/Dune/1) and the local Research and Development department have reviewed and approved this study. You can thus be assured that your time and contribution to this study will be worthwhile and will contribute towards the body of knowledge in this important area, improvement of clinical processes, NHS services and patient outcomes.

#### **Confidentiality**

All reasonable means will be taken to protect the anonymity of participating individuals and no person identifiable data will be used in this study. You will be given an opportunity to decide whether your hospital or technology is identified in the study and all efforts will be made to ensure anonymity where requested. However, it is important for you to understand that some healthcare technologies are specialist in nature and are used in only a few NHS organisations. As such other NHS Trusts or wider healthcare organisations may be able to link the study to your hospital.

Raw data generated in the course of this study will be handled in accordance with the Data Protection Act 1998 and The Open University's research code of practice. Transcribed electronic data will be kept on The Open University computer, where only the researcher and supervisors will have access using secure passwords. All data used in the thesis will be anonymised following collection and on completion the thesis will be deposited into The Open University theses repository. Audio recordings will be retained for a period of twelve months following completion of the project before being deleted. However, audio-recorded data can be deleted immediately after the interview is summarised at your request. The researcher has an ethical duty to report to the relevant authority any illegal, dishonest or unethical practice that he might become aware of during the study. The researcher is covered by The Open University's professional indemnity insurance while undertaking this study.

## What will happen to the results of the research study?

Results of the study will be reported in a PhD thesis at the end of 2014 and other published outlets such as conferences, peer group forums and journal publications. All participants will be provided with links to the final PhD thesis report and other published outputs.

## Research funding

The Open University funds this research project as a PhD studentship under the supervision of Professor Joyce Fortune and Dr Clive Savory from the Communication and Systems Department.

## Contact for further information:

Richard Dune, PhD Student  
Communication and Systems Department  
Faculty of Mathematics, Computing and Technology  
The Open University | Walton Hall | Milton Keynes | MK7 6AA  
Tel: (0) 1908 653964

## Contact for Comments and issues related to this research:

Dr. Nicky Moss, Head of Department  
Communication and Systems Department  
Faculty of Mathematics, Computing and Technology  
The Open University | Walton Hall | Milton Keynes | MK7 6AA  
Tel: +44 (0) 1908 654933  
Email: [n.g.moss@open.ac.uk](mailto:n.g.moss@open.ac.uk)



Faculty of Mathematics, Computing and Technology

Department of Communication and Systems

The Open University  
Walton Hall, Milton Keynes  
United Kingdom  
MK7 6AA

Tel +44 (0) 1908 653964  
Fax +44 (0) 1908 653718

Email: richard.dune@open.ac.uk

Consent Form Version 2 - 22/11/2012

The Open University HREC Approval Ref: HREC/2012/#1213/Dune/1

**Study title – Evaluation of Clinical Decision-Support Technology in the NHS**

I confirm that I have read and understood the participant information sheet and I have had an opportunity to ask questions about my participation	
I understand that my participation in this study is entirely voluntary and that I am free to withdraw at any time before the results of the study are published	
In the event that I withdraw from the study, I agree that my data collected up to that point can be used in the study and related published outputs	
I agree to the interview being audio recorded and that anonymised data collected will be used in research publications	
I agree that my data gathered in this study will be kept securely at The Open University and that the thesis produced from this study will be added to The Open University repository. Electronic data will be kept securely in The Open University server	
I agree to take part in this study	

Name of Participant

**Date and time**

**Signature**

Name of Researcher

**Date and time**

**Signature**

Should you require further information before, during and after the study, please contact the Chief Investigator using the details below:

Richard Dune, PhD Student

Communication and Systems Department | Faculty of Mathematics, Computing and Technology

The Open University | Walton Hall | Milton Keynes | MK7 6AA

Tel: 01908 653964

Mob: 07824323573

Email: richard.dune@open.ac.uk

## Appendix 6 – HREC Ethics Approval



The Open University

From Dr Duncan Banks  
Chair, The Open University Human Research Ethics Committee  
Email d.banks@open.ac.uk  
Extension 59198  
To Richard Dune, Mathematics, Computing and Technology  
Subject "Clinical decision support systems (CDSS) evaluation in the NHS."  
Ref HREC/2012/#1203/Dune/1  
Red form  
Submitted 28 June 2012  
Date 10 July 2012

### Memorandum

This memorandum is to confirm that the research protocol for the above-named research project, as submitted for ethics review, **approved** by the Open University Human Research Ethics Committee.

Please make sure that any question(s) relating to your application and approval are sent to [Research-REC-Review@open.ac.uk](mailto:Research-REC-Review@open.ac.uk) quoting the HREC reference number **HREC/2012/#1203/Dune/1**. We will endeavour to respond as quickly as possible so that your research is not delayed in any way.

At the conclusion of your project, by the date that you stated in your application, the Committee would like to receive a summary report on the progress of this project, any ethical issues that have arisen and how they have been dealt with.

Regards,

Dr Duncan Banks  
Chair OU HREC



From Dr Duncan Banks  
Chair, The Open University Human Research Ethics Committee  
Email duncan.banks@open.ac.uk<sup>[1]</sup>  
Extension 59198  
To Richard Dune, Mathematics, Computing and Technology  
Subject "Evaluation of Clinical Decision-Support Technology in the NHS."  
Ref HREC/2012/1213/Dune/2  
Red form  
Submitted 5 December 2012  
Date 6 December 2012

## Memorandum

This memorandum is to confirm that the **modification to the** research protocol for the above-named research project, as submitted for ethics review, is **approved** by the Open University Human Research Ethics Committee.

Please make sure that any question(s) relating to your application and approval are sent to [Research-REC-Review@open.ac.uk](mailto:Research-REC-Review@open.ac.uk) quoting the HREC reference number above. We will endeavour to respond as quickly as possible so that your research is not delayed in any way.

At the conclusion of your project, by the date that you stated in your application, the Committee would like to receive a summary report on the progress of this project, any ethical issues that have arisen and how they have been dealt with.

Regards,

Dr Duncan Banks  
Chair OU HREC

<sup>[1]</sup> please note the change in email address

## **Appendix 7 - GfREC Guidance – Main Changes to the REMIT of RECs and study Trust approval**

The following types of research do not normally require review by an NHS REC within the UK Health Departments' Research Ethics Service. Projects being undertaken as part of an educational qualification however, still require review from a University REC.

### **1) Research involving staff**

Research involving NHS staff recruited as research participants by virtue of their professional role, is excluded from the normal remit of RECs under the harmonised edition of GfREC.

### **2) Research involving previously collected, non-identifiable information**

Research involving previously collected, non-identifiable tissue samples in accordance with the terms of donor consent is generally excluded from REC review.

However, REC review would be required if any of the following applied:

- a) Consent for research has not been given, or the research is not within the terms of the consent.
- b) The research also involves removal, storage or use of new samples from the living or deceased.
- c) The research also involves the use of identifiable information held with the samples.

### **3) Research involving acellular material**

Research limited to use of human biological material not consisting of or including cells (e.g. plasma, serum, DNA) is also generally excluded from REC review.

However, REC review would be required if the research involved:

- a) Prospective collection of tissue samples from patients in order to extract acellular material for research.
- b) Prospective collection of information from patients.
- c) Use of previously collected information from which patients could be identified by the researchers.
- d) Analysis of DNA in material from the living, where consent for research is not in place from the person whose body manufactured the DNA.

### **4) Research involving previously collected, non-identifiable information**

Research involving previously collected, non-identifiable information. This exception also

applies to research undertaken by staff within a care team using information previously collected in the course of care for their own patients or clients, provided that data is anonymised in conducting the research.

However, REC review is required for research involving:

- a) Prospective collection of information from patients or service users for research.
- b) Use of previously collected information from which patients or service users could be identified by researchers outside of the usual care team (either directly from that information, or in combination with other information, or likely to come into possession, their possession).

#### **5) Research involving premises and facilities**

Research limited to the use of or access to NHS premises and facilities no longer REC review, provided that review is not required under other provisions of GAFREC.



Dear Mr Dune,

**Study Title: Evaluation of Clinical Decision-Support Technology in the NHS**

Thank you for submitting the above study for consideration by the Research & Development Office. I am pleased to inform you that your study has been approved.

The documents approved for use in this study are:

Document	Version	Date
Protocol		Dec 12
Participant Information Sheet	2	30.11.2012
Consent Form	2	30.11.2012

- Should you wish to make any changes to the documents listed above, you must obtain R&D approval prior to use.
- An Annual Progress Report (APR) should be submitted to the main research ethics committee (REC) once a year throughout the trial or on request by R&D. The first report is due on **18<sup>th</sup> December 2013**. In addition, for CTIMP studies, a Development Safety Update Report (DSUR) should be submitted to the MHRA and the REC once a year. Guidance on the DSUR can be found in SOP 41 'Preparation and Submission of Annual Progress Reports and Development Safety Update Reports'.
- Notification of any serious breaches of GCP or the trial protocol must be reported to the R&D Department and a DATIX Clinical Adverse Event form completed within 24 hours of any suspected breach being identified and confirmed.

R&D Reference: RD115512

Version 4, 01.11.2012

Page 1 of 2

Your research sponsorship & Indemnity is provided by The Open University.

Your project may be subject to ad hoc audit by our department to ensure these standards are being met.

May I take this opportunity to remind you that, as a researcher, you must ensure that your research is conducted in a way that protects the dignity, rights, safety and well-being of participants. Trust R&D Approval assumes that you have read and understand the Research Governance Framework and accept that your responsibilities as a researcher are to comply with it, the Data Protection and Health & Safety Acts.

The Trust wishes you every success with your project.

Yours sincerely

## Appendix 8 - Coding practices (examples)

Our Strategic Health Authority were producing report where they were ranking high... [coded as NHS rankings] it was interesting and slightly frustrating you start seeing some Trust whereby you know their processes start reporting 98% [CQUIN compliance] success rates and when you talk to them, you realise they don't count day surgery and don't consider such patients as in-patients... [coded as self-reporting problems] that's actually not the spirit of what NICE want. ... It seems to me no one is looking hard at the figures.

T1LeadConsultant

I am under no illusion if it wasn't for the CQUIN money I don't think this [successful T1 implementation] could have happened [coded as CQUIN incentives] and it's absolutely pivotal to have the support of the executives and that really helped as it started the ball rolling... once they signed up to it, finance became less of an issue... [coded as importance of executive support]

T1LeadConsultant

There wasn't time for a pilot... it [T1] had to go [live]... It did work but there were departments where it absolutely caused mayhem... [coded Trustwide implementation chaotic] the way it was worded meant that every admission had to be risk assessed, even minor procedures... for example, in the Surgical Day Unit, day surgery lists for minor procedures were held back... on the day the tool went live, most of the VTE project team were away and I had to deal with the various issues... [coded as unanticipated implementation challenges]

T1ImplementationNurse

I don't think it was effective (online VTE training for nurses)... [coded as ineffective training] it was almost like a task...I remember it being quite hard to work through, not

that it was a difficult assessment but not very user friendly [coded as training not user friendly]. Those online assessments are sometimes like tick box exercises... how much that actually adds to the assessment being complied with, I really don't know... I don't think it adds much knowledge as to why it's important. I don't think the training itself added much to my knowledge... I can't even remember what was in that training... it just doesn't stick really [coded as training perceived insufficient].

T1ANP

Some GPs have the skills to do stable prostate cancer follow ups and some are less comfortable... [coded as variable GP skills]. As GPs are organised in clusters, it is challenging to agree on the logistical and organisational issues... who is going to head up the service in that area?; where will the data reside?; and how to filter patients into that particular service... [coded as organisational and logistical problems] the successful ones we have had are where the organisation is acting as a hub and there is a structure in place where patients actually merge to one particular [agreed] location, and they might have several patients from various outlying GPs... that problem has been solved already before we got there so that's good, but more complicated in some situations... [coded as implementation challenges]

T2Developer

Making one computer talk to another is relatively straightforward. The problem is not technical... it is dealing with the hospitals in general, protocols, the red tape, getting to the data... [coded as organisational challenges]. I understand the many good reasons for data security... The biggest problem is securing the hospital IT resource, obtaining information governance clearance, to release information or make sure that information will be used in a safe manner... [coded as information governance]

concerns] agreeing with IT departments a method that is mutually acceptable, especially across Trusts because we are having to get two discreet teams jelling together... it's quite a challenge [coded as co-working challenges]

T2Developer

You couldn't rely on the system [T2] to generate all the answers for you or just accept everything it recommends... [coded as CDSS not panacea] it will generate a lot of medical issues but sometimes you need to have the clinical knowledge to deal with the practical issues... clinical knowledge such as the treatments they have had and their side effects, reading the results and appropriate follow up... [coded as clinical knowledge and experience essential]

T2Nurse1

They [T2 developing Trust and their commercial partners] were looking for site to roll it out (T2)...we have previously collaborated with them [T2 developing Trust]... we knew the guys who developed it [coded as prior knowledge of developers] and it seemed to be working well for them... they had a similar unit to ours, so we thought we would give it a go particularly as the implementation costs were going to be paid by their commercial partners... [coded as financial incentives]

T2LeadConsultant

The best evaluation is comparing Consultants' decisions and what the computer would have done but that takes too much time [coded as RCT gold standard]. The attraction for us is that they [T2 developing Trust] had already done that so we were happy to take their word and we were only happy to take their word because we knew them from before and we trusted them, whereas if it was a completely commercial outfit

that would make us more suspicious as to what they were trying to sell to us... [coded as trust in peer evaluations]

#### T2LeadConsultant

It was a big job [the implementation]... quite a lot of nurses didn't even have computer login details or access to the system for electronic charts... lots of facilitation, getting line managers to sign things off, just trying to make things easy for them... account set up [problems with] were unexpected, although there was an inkling... [coded as unexpected organisation and logistical issues] little has changed regardless of the number of projects that have been rolled out since... we are still bringing forms to get people logged in... it's becoming role specific, nurses have access, but there is a huge gap with healthcare assistants, but now with ESR, it's becoming more important for them to use computers... [coded as computer apathy]

#### T3LeadNurse

Our approach to user training was flexible and pretty mixed... It was easier to fit training around the nurses' station rather than sisters' office, so the nurses were still accessible on the ward if [they were] needed [coded as flexible training]. It was broken training, because emergency buzzers, phones ringing... but they [the nurses] had faith in that we would let them attend to their work, and they actually concentrated during training... We also went on observation rounds, and were all in nurses' uniforms. We were able to freely go around the ward, be with patients and looked professional... it

helps when patients know that you are a nurse... [coded as trust in implementation team]

#### T3LeadNurse

It's a bit like big brother watching you... [coded as negative CDSS perception] to keep an eye on people who don't know what they are doing, in a helpful way. It's like an idiot's guide when to do observations [coded as CDSS useful to some staff]. It takes away professional judgment and tells you what and when to do it... [coded as CDSS deskills nurses]. It's for the Trust and everyone must have it... legally they (senior Trust managers) can prove that patients' observations have been done because they have it down on a computerised system [Trust 'covering back']. I am not dead against it... we got used to it and I probably wouldn't want to go back to paper charts now...The idea was that doctors aren't called unnecessarily, and they can look up for observations without having to go to the ward. We do that ourselves when we get referrals from other wards [general wards]... [coded as CDSS has benefits]

#### WardSister1

It was an agile approach to software development, going back and forth and trying to support the rollout [coded as agile CDSS development]. It was their [T3 supplier] first ever roll out of such magnitude so they were quite willing to work with us and provide the functionalities that were needed... [coded as CDSS developers supportive] It had benefits for them because if we were asking for something, other Trusts were bound to want the same... they actually began to have a comprehensive package that they could sell . . . they wanted us to roll it out Trust wide and could then say, "we have this

system, which we rolled out in this big university hospital in 6 months..." [coded as commercial benefits for CDSS developers]

T3LeadNurse

Night sisters and the Outreach team can now identify them (deteriorating patients) and intervene appropriately... [coded as CDSS benefits] for a system that I imagine costs a lot of money, I'm not sure whether it's worth it. If it has cost more than making sure that you have enough staff on the wards, senior nurses to keep an eye on things rather than machines on the ward ... you can't beat someone being physically there... [coded as money could be better spent elsewhere] it is taking away people's ability to make decisions and in the long term may deskill people... [coded as CDSS deskill nurses]

WardSister1

Contextual factors affecting evaluation	Key questions	T1	T2	T3
Existing electronic technology infrastructure				
<i>Circumstances necessitating CDSS adoption</i>	Did CDSS replace paper-based charts?	Yes	Partially	Yes
<i>Regulatory basis of CDSS adoption</i>	Was CDSS based on NICE guidelines?	Yes	Yes	Yes
<i>Other electronic systems already in use</i>	Were there legacy systems used for the clinical problem?	No	Yes	No
<i>Pressures associated with guideline adherence</i>	Were there penalties for non-adherence with guidelines?	Yes	No	Yes
Organisational readiness for evaluation				
<i>CDSS effects on existing care systems</i>	Did the CDSS lead to changes in working patterns and care delivery?	Yes	Yes	Yes
<i>Assessment of existing systems</i>	Did project teams systematically assess existing legacy systems?	Yes	No	No
<i>Intended users' perception of CDSS</i>	Was CDSS contested by intended users?	Yes	Yes	Yes
<i>Responsibility for CDSS maintenance</i>	Was CDSS centrally managed by the ICT department?	Yes	No	Yes
Attitudes towards electronic technology				
<i>Clinicians' attitudes towards legacy systems</i>	Did clinicians use legacy systems prior to CDSS adoption?	Yes	Yes	Yes
<i>Nurses' attitudes towards legacy systems</i>	Did nurses routinely use legacy systems prior to CDSS adoption?	No	Yes	No
<i>Extend of technology usage prior to CDSS adoption</i>	Were clinical processes technology enabled prior to CDSS adoption?	No	Yes	Partially
<i>Perceived CDSS usefulness by nurses</i>	Did nurses consider CDSS useful for clinical decision making?	No	Partially	Partially
<i>Perceived CDSS usefulness by clinicians</i>	Did clinicians consider CDSS useful for clinical decision making?	No	Yes	Yes
<i>Perceived CDSS usefulness by managers</i>	Did managers consider CDSS useful for clinical decision making?	Yes	Yes	Yes
<i>Perception that CDSS is under-utilised</i>	Did key stakeholders consider CDSS to be underutilised?	Yes	Yes	Yes
External technology context				
<i>Regulatory recommendations to adopt technology</i>	Were there regulatory recommendations to adopt CDSS?	Yes	Yes	Yes
<i>Recommendations from professional bodies</i>	Was CDSS recommended by professional/collegiate bodies?	Yes	Yes	Yes
<i>Influence from peer groups</i>	Was there influence from peer groups to adopt CDSS?	Yes	Yes	Yes
<i>Commercially available systems</i>	Were there any commercially available systems at time of adoption?	No	No	No
<i>Availability of "home-grown" systems</i>	Were there any 'home grown' systems available at time of adoption?	No	Yes	Yes

Table 10.1 Key contextual factors affecting evaluation

Key evaluation issues	Comments	T1	T2	T3
Owners of evaluations				
<i>Project Team leaders</i>	Did the Project Team leaders own the evaluations?	Yes	No	Yes
<i>Trust Board</i>	Did the Trust Board leaders own the evaluations?	Yes	No	Yes
<i>External evaluators</i>	Did external evaluators leaders own the evaluations?	Yes	Yes	Yes
Customers of evaluations				
<i>Trust Board</i>	Were the Trust Board customers of evaluation?	Yes	No	Yes
<i>Regulatory authorities</i>	Were the regulatory customers of evaluation?	Yes	Yes	Yes



<i>Commissioners</i>	Were the Commissioners customers of evaluation?	Yes	Yes	Yes
<i>Barriers to evaluations</i>				
<i>Lack of funding</i>	Was lack of funding highlighted as a barrier to evaluation?	Yes	Yes	Yes
<i>Priority for evaluation given to guideline adherence</i>	Was evaluation priority given to guideline adherence?	Yes	Yes	Yes
<i>Assumption that CDSS improved outcomes</i>	Were there assumptions that CDSS had improved outcomes?	Yes	Yes	Yes
<i>Lack of support from commissioners</i>	Was lack of support from commissioners highlighted as an issue?	No	Yes	No
<i>Concern about evaluation related ethical issues</i>	Were there concerns about evaluation related ethical issues?	Yes	No	No
<i>Conflicting evaluation interests</i>	Were there conflicting evaluation interests between key stakeholders?	Yes	No	Yes
<i>Purposes of evaluation</i>				
<i>Compliance with guidelines</i>	Were evaluations carried out to assess guideline compliance?	Yes	Yes	Yes
<i>Improving clinical practice</i>	Were evaluation carried out to improve clinical practice?	Yes	Yes	Yes
<i>Improving patient outcomes</i>	Were evaluation carried out to improve patient outcomes?	Yes	Yes	Yes
<i>Technology investment justification</i>	Were evaluation carried out to justify technology investment?	No	No	Yes
<i>Evaluation inputs</i>				
<i>Internal funding for evaluation</i>	Was there internal funding for evaluations?	Yes	No	Yes
<i>External funding for evaluation</i>	Was there external funding for evaluations?	No	Yes	
<i>Secondment of key personnel</i>	Were key personnel seconded for CDSS project and evaluation?	Yes	No	Yes
<i>Use of existing administrative systems</i>	Were existing administrative systems used for evaluations?	Yes	No	Yes
<i>Evaluations undertaken</i>				
<i>Technical efficacy testing during development</i>	Did evaluators carry out technical efficacy tests during CDSS development?	Yes	Yes	Yes
<i>Clinical efficacy testing during development</i>	Did evaluators carry out clinical efficacy tests during CDSS development?	Yes	Yes	Yes
<i>User testing during development</i>	Did evaluators carry out user tests during CDSS development?	Yes	Yes	No
<i>Validation of paper algorithms during development</i>	Were paper algorithms validated during CDSS development?	Yes	Yes	No
<i>Validation of software algorithms during development</i>	Were software algorithms validated during CDSS development?	Yes	Yes	No
<i>Internal cost effectiveness audits</i>	Were internal cost effectiveness audits carried out?	Yes	No	Yes
<i>Internal patient safety audits</i>	Were patient safety audits carried out?	Yes	No	
<i>External clinical effectiveness audits</i>	Were clinical effectiveness audits carried out?	Yes	No	Yes
<i>Benefits realisation studies</i>	Were benefits realisation studies carried out?	No	No	Yes
<i>Service evaluations</i>	Were service evaluations carried out?	Yes	No	Yes
<i>Informal evaluations</i>	Were informal evaluations carried out?	Yes	Yes	Yes
<i>Root cause analyses</i>	Were root cause analyses carried out?	Yes	No	No
<i>Documentation of internal evaluations</i>	Was there documentation for internal evaluations?	Yes	No	No
<i>Cost effectiveness studies</i>	Were cost effectiveness studies carried out?	No	No	Yes
<i>Patient survey</i>	Were patient survey's completed?	No	Yes	No
<i>Documentation of external evaluations</i>	Were external, evaluations documented?	No	No	Yes
<i>Interactions between evaluators and CDSS users</i>				
<i>Users' inadvertent responses to evaluators</i>	Did users inadvertently respond to evaluators?	Yes	Yes	Yes
<i>Assessment of user and CDSS interaction</i>	Did evaluators assess user and CDSS interaction?	No	No	No
<i>Evaluation information fed back to users</i>	Was evaluation information fed back to users?	Yes	No	No

Evaluation outcomes		Yes	Yes	Yes
<i>Compliance with national guidelines</i>	Did evaluations show compliance with national guidelines?	Yes	Yes	Yes
<i>Meeting national targets</i>	Were evaluations used to show adherence with national targets?	Yes		Yes
<i>Continued project funding</i>	Was CDSS evaluation used to justify continued CDSS funding?	Yes	Yes	Yes
<i>Additional departmental funding</i>	Was CDSS evaluation used to justify continued departmental funding?	Yes		Yes
<i>Perceived improvement in clinical practice</i>	Were CDSS considered to have improved clinical practice?	Yes	Yes	Yes
<i>Perceived improvement in patient outcomes</i>	Were CDSS considered to have improved patient outcomes?	Yes	Yes	Yes

Table 10.2 Cross-case analysis using CDSS evaluation framework